Information Flow in Linear Systems

Se Yong Park

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2013-233
http://www.eecs.berkeley.edu/Pubs/TechRpts/2013/EECS-2013-233.html

December 20, 2013

Information Flow in Linear Systems

by

Se Yong Park

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Electrical Engineering and Computer Sciences

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:
Associate Professor Anant Sahai, Chair
Professor Venkat Anantharam
Professor Andrew Packard
Professor David Tse

Fall 2013

Information Flow in Linear Systems

Abstract

Information Flow in Linear Systems

by

Se Yong Park
Doctor of Philosophy in Electrical Engineering and Computer Sciences

University of California, Berkeley

Associate Professor Anant Sahai, Chair

Modern control systems have their own unique features that distinguish them from classical control systems. In many cases, there inherently exist multiple distributed controllers and so communication networks can be introduced to connect them. Due to these features, it is challenging to design efficient controller and transceivers for modern control systems. Practically, we must answer questions like "How reliable do we need the communication channels to be to achieve the desired control performance?", "What information should be exchanged between controllers?" and "What are the optimal controller and transceiver structures?" All these practical equations are related to one theoretical question "Can we understand the information flows between controllers?" In other words, the controllers communicate with each other explicitly through the communication networks and also implicitly through the plants, and we have to understand this information flow for control. In this thesis, we consider three seemingly simple but fundamental problems to understand explicit and implicit information flows for control, as initial building blocks for a theory that we hope will eventually lead to novel and efficient designs for modern control systems.

In the first technical chapter, we consider Kalman filtering problems when the observations are intermittently erased or lost. Practically, this problem is the simplest model for the packet losses that can happen in communication networks connecting distributed controllers. Theoretically, by relating the erasure probability of the channel with the stability of the control system, we can measure the minimum quality requirements for uncoded information that has to flow to stabilize the system. It was known that the Kalman filtering estimates are mean-square unstable when the erasure probability is larger than some critical value, and stable otherwise. But what that critical value actually is has been open for years. Unlike prior work that tried to connect with Lyapunov stability, we connect with observability to completely characterize the critical erasure probability. We introduce a new concept of *eigenvalue cycles* which captures the periodicity of systems, and characterize the critical erasure probability based on this. It is also proved that eigenvalue cycles can be easily broken if the original physical system is considered to be continuous-time — randomly-

dithered nonuniform sampling of the observations makes the critical erasure probability almost surely $\frac{1}{|\lambda_{max}|^2}$, the best that could be hoped for even with arbitrarily complex coding. This implies that the rank of the observability gramian can be thought of as the amount of information that the estimator learns about linear systems, and nonuniform sampling helps maximize that rank. Furthermore, different subspaces of the states can be thought of as the source of information flows and separated as long as they belong to different eigenvalue cycles.

In the second technical chapter, to understand implicit information flows we consider distributed linear systems without explicit communication networks. To do this, we build a unified view of both network coding and decentralized control. Precisely speaking, we consider both as linear time-invariant systems by appropriately restricting channels and coding schemes of network coding to be linear time-invariant, and the plant and controllers of decentralized control to be linear time-invariant as well. First, we apply linear system theory to network coding. We introduce a novel technique that we call *Network Linearization.* This technique gives a way of converting an arbitrary relay network to an equivalent acyclic single-hop relay network. Based on network linearization, we prove that the fundamental design limit, mincut, is achievable by a linear time-invariant network-coding scheme regardless of the network topology. Unlike previous approaches relying on graph theory, we use linear system theory and linear algebra, and exploit the fact that there can be multiple network representations for a given algebraic transfer function. For broadcast and unicast problems, unintended messages at receivers turn out to be modeled as secrecy constraints after network linearization.

Having built a linear-systems view of network coding, we turn it around to view decentralized linear control systems. We argue that linear time-invariant controllers in a decentralized linear system "communicate" via linear network coding to stabilize the plant. To justify this claim, we revisit classical stabilizability results concerning fixed modes. We give an algorithm to "externalize" the implicit communication between the controllers that we believe must be occurring to stabilize the plant. Based on this, we show that the stabilizability condition for decentralized linear systems comes from an underlying communication limit, which can be described by the algebraic mincut-maxflow theorem. With this re-interpretation in hand, we also consider *stabilizability over LTI networks* to emphasize the connection with network coding. These results confirm the intuition that there are implicit information flows in distributed control systems which we can visualize. Moreover, the rank of subspaces are the proper measure of information for linear systems when we consider stabilizability.

In the third and fourth technical chapters, we go beyond stabilizability and study how the size of implicit information flows constrain the optimal control performance. To do this, we must allow arbitrary controllers without imposing linearity constraints. In particular, we focus on scalar decentralized average-cost infinite-horizon LQG problems with two controllers. For fast-dynamics systems — when the eigenvalue of the system is large —, it is shown that the best linear controllers' performance can be an arbitrary factor worse than the optimal nonlinear controller performance.

To understand the required nonlinearity in such control systems, we caricature bit-levels of the states as different subspaces, and the rank of those subspaces as the amount of information. In other words, we take a *linear view of nonlinearity.* Based on this insight, we propose a simple set of finite-dimensional nonlinear controllers, and prove that the proposed set contains easy-to-find approximately optimal strategies that achieve within a constant ratio of the optimal quadratic cost. The insight for the nonlinear strategies comes from revealing the relationship between implicit information flow in control and wireless information flow. More precisely, we discuss a close relationship between the high-SNR limit in wireless communication and the fast-dynamics case in decentralized control, and justify how the proposed nonlinear control strategy can be understood as exploiting a kind of generalized degree-of-freedom gain in wireless communication theory. For a rigorous justification of this argument, we develop new mathematical tools and ideas. We extend Witsenhausen's counterexample to MIMO (multiple-input multiple-output) Witsenhausen's counterexamples, just as wireless communication extends the scalar AWGN (additive white Gaussian noise) channel to MIMO channels and from there, eventually tackles multi-terminal problems. To reveal the relationship between infinite-horizon problems and generalized MIMO Witsenhausen's counterexamples, we introduce the idea of *geometric slicing* that plays a role like that of cut-set bounds in communication theory. To analyze nonlinear strategy performance, we introduce an approximate-comb-lattice model for the relevant random variables. For the slow-dynamics cases — when the eigenvalue of the system is small —, we prove that single-controller optimal strategies —linear strategies— are constant-ratio optimal among all distributed control strategies.

Understanding the nature of information flow for control should eventually lead to a unified theory for control and communication. We believe the parallelism between control information flow and wireless information flow is not just a coincidence but strong evidence for such a unified theory. However, still lots of concepts and ideas in control and communication remain separate and have not been related — for example, secrecy, interference alignment, and scaling laws in communication. Therefore, further research is required to continue uncovering the fundamental relationship between control and communication. Moreover, we also have to think how to leverage this understanding in practical system designs, and how to build efficient distributed controllers and transceivers for modern control systems.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

A goal of both control and communication theory is building efficient systems. Control theory studies how to achieve the best performance with minimum effort. Communication theory studies how to convey information with minimum cost. In other words, both problems can in principle be formulated as optimization problems, and communication problems may not have to be distinguished from control problems. In fact, until the late 1940s, the concept of cybernetics [106] had not distinguished control and communication as separate fields.

However, Shannon's revolution [93] overturned this classical paradigm. Relying on the fact that communication systems can switch much faster than human recognition, Shannon conceived block-coding strategies. It was well-known that as the length of i.i.d. random variable sequences gets longer, the empirical distribution converges to the probabilistic distribution. However, it was only after Shannon's novel application of this fact that the concept of entropy was discovered as a *useful* measure of information (it had been an abstract and philosophical concept before). Since the discovery of entropy, communication theory —information theory in a broader sense— has separated from control theory and become an independent research area.

While information theory has been developing mathematical tools to quantify a philosophical concept of information, control theory has mainly focused on extending the classical optimization framework and finding practical applications. However, intuitively information and control have to be deeply connected, since to control a plant we first need information about the plant. This intuitive connection motivated both information and control theorists to explore the relationships between the two theories [109, 45, 89, 8, 97, 26, 67, 86, 37].

More importantly, a new concept of modern cyber-physical systems [96, 41, 57], which have both control and communication parts in them, recently emerged. To properly understand these new systems, the unification (at the very least, partial compatibility) of control and communication theory is becoming crucial. Modern systems differ from classical control systems in two aspects. The first difference is the inherent distributedness of the systems. Unlike classical centralized control systems,

Figure 1.1: An Artist's Picture of Intelligent Transportation Systems [47]

as system size scales, it is becoming physically impossible to introduce centralized controllers. The second difference is that since controllers are naturally distributed, wired/wireless communication technology can be used to connect the distributed controllers.

Figure 1.1 shows an artist's picture of intelligent transportation systems as an example of modern systems. Until the last century, each vehicle had been considered as a separate system, and controlled by one dedicated centralized controller. However, as the number of vehicles scales, each vehicle starts to interact with (e.g. potentially collide with) each other. One goal of intelligent transportation systems is to design controllers which avoid negative interactions (e.g. avoid collisions) between the vehicles. To design such systems from a control-theoretic point of view, we can no longer consider different vehicles as separate systems. We have to model all the vehicles collectively as one big plant. The individual controllers dedicated to different vehicles should be thought of as distributed controllers with partial information (around the corresponding vehicle) and partial control (over the corresponding vehicle) of the big plant. Furthermore, we can employ current wireless communication technologies to connect these distributed controllers to share their information as well as having them using sensors to view each others' actions.

However, control theory and the mathematical tools that currently exist are not enough to decisively address core practical engineering questions in the design of these modern systems. For example,

- How much wireless spectrum do we have to allocate for communication?
- How can we guarantee the stability of systems?
- How should we process communication and control signals?
- What kind of architecture should be used for controllers, and for communication for control?

Furthermore, the distributed nature of systems and interwoven communication networks are also found in other modern systems including the smart grids [51], manufacturing [13], civil infrastructure [111], and health care [58].

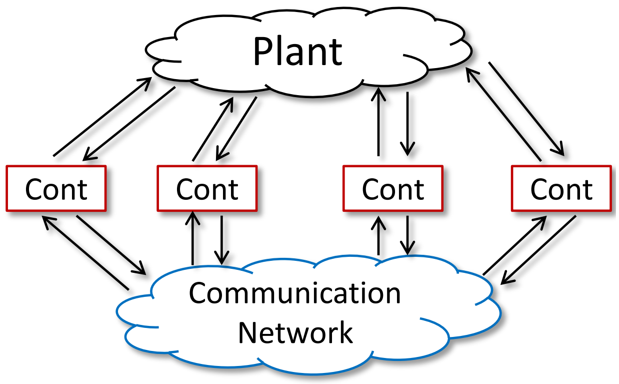


Figure 1.2: A Conceptual Diagram of Modern Cyber-Physical Systems

Therefore, to answer such fundamental questions about modern control systems, we have to build a new theory that embraces both control and communication theory. Modern systems can be represented by the conceptual diagram of Figure 1.2. In Figure 1.2, we can see multiple controllers with partial information and control interacting with one big plant which models whole physical systems. One big communication network which includes all individual channels/links between the controllers is also shown in the diagram.

The main difference between classical centralized and modern distributed systems is information flows. In classical centralized systems, the conceptual information flow for control is simple enough to be ignored. Information (or uncertainty) is generated at the plant, flows to the controller as the controller observes the plant, and dissipates as the controller controls the plant (and thereby removes uncertainty at the plant) [67]. Furthermore, within the controller, classical control theory assumes every component is connected with infinite capacity, perfectly reliable, and zero-delay links. Therefore, there was no need to even measure the information flow required to control systems.

However, this is not the case for modern distributed systems. In modern control systems, controllers consist of lots of different components, which might be connected by shared communication buses with bounded-capacity or unreliable communication channels. Therefore, each controller has only partial information about the plant, and they may want to communicate with each other to reduce the uncertainty that each faces. This communication can be done both explicitly through the communication network and implicitly through the plant — this point will be clarified in Chapter 3. In other words, the information generated at the plant is distributively observed and controlled by multiple controllers, and flows through different controllers until it is actually dissipated by the control. As a result, even finding a conceptual flow diagram corresponding to a specific uncertainty is challenging. Measuring information flows for control, and thereby characterizing the tradeoff between control performance and channel quality becomes practically important.

In short, to understand modern distributed systems from a control-theoretic point of view,

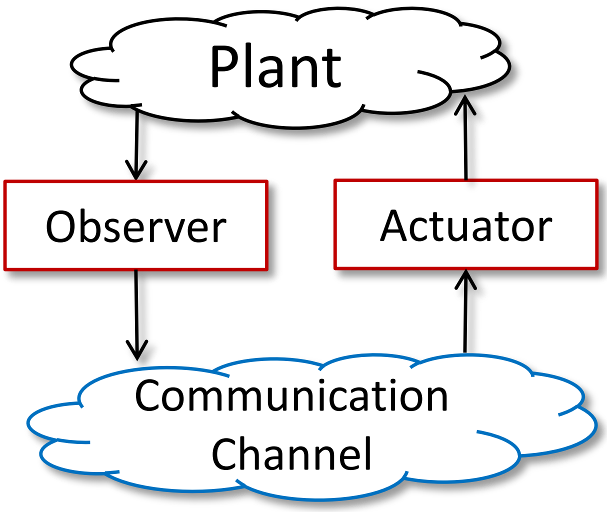


Figure 1.3: A Conceptual Diagram of Control-Over-Communication-Channel Systems

we have to understand information flows for control and develop mathematical tools to measure them. To further this end, we take a bottom-up approach. Since modern control problems are too complicated to tackle directly, we will consider simpler but canonical problems which capture some essential aspects of the original problems. By doing this, we expect to find the fundamental nature of the problems in minimal form.

The first simpler problem is control over a communication channel [97]. As shown in Figure 1.3, there are only two controllers in this problem. One controller, which we call the observer, can only observe the plant, and the other controller, which we call the actuator, can only act on the plant. Therefore, to control the system, the observer has to communicate information to the actuator. Since the explicit communication channel is the only medium which connects these two, the information for control has to flow through this explicit communication channel. By relating the reliability of the communication channel with the control performance, we can expect to find a proper measure of information for control [87].

The second simpler problem is distributed control *without* communication networks, as shown in Figure 1.4. Of course, modern cyber-physical systems are equipped with communication networks. However, to make the best use of the communication networks, we have to decide what information to send. Understanding control information flow without explicit communication networks can be greatly helpful in making such decisions by providing a baseline against which improvement can be evaluated. After all, systems with communication networks are only more complicated than systems without communication networks.

The challenge is that when there are no communication networks, information for control must flow implicitly through the plant. In other words, controllers can embed information in their

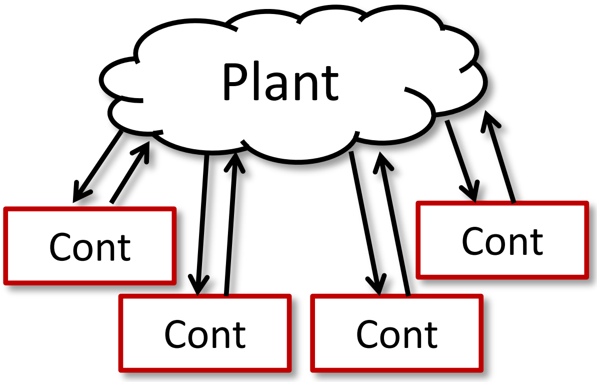


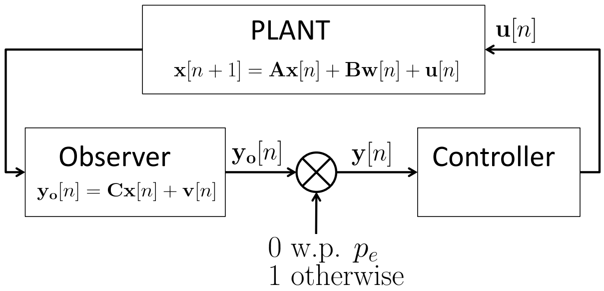
Figure 1.4: A Conceptual Diagram of Distributed Systems without Communication Networks

actions. Then, the information will be temporarily stored in the plant until another controller receives this information by its observations. Therefore, by studying this problem, we can understand how information flows in systems with multiple controllers.

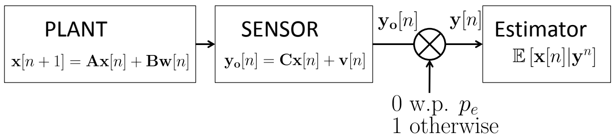
However, even though the concept of information flow for control sounds simple and the diagrams of Figure 1.3 and 1.4 look simple, getting a mathematical or even qualitative formulation of the concept is challenging. First of all, the classical notions of entropy and Shannon capacity are not enough to measure the amount of information for control [86]. One crucial assumption in deriving entropy and capacity for communication systems is the assumption that delay is not important. Electromagnetic waves propagate at the speed of light and VLSI chips process signals at giga hertz. Therefore, when the source or the destination of information is a human whose recognition cycle is much slower than chip speeds, we can tolerate long delay assumptions. However, in modern cyber-physical control systems, the sources and destinations are physical systems which can operate much faster than humans and evolve over time [98]. Furthermore, as the system scales, more and more controllers share a common communication network. Communication delay could increase significantly due to congestion, and information could even be lost in the communication network. Therefore, we have to define a proper notion of information that takes into account delay and unreliability in communication and the dynamics of plants.

Second, when there exist multiple controllers, finding information flow paths can be challenging. The conceptual information path in the system of Figure 1.3 looks obvious since one controller can only observe and the other controller can only act. However, when there are multiple controllers which can both observe and act on the plant like the system of Figure 1.4, finding a conceptual information path is not trivial. It is not clear what the sources, relays and destinations of information are. Without a conceptual picture of information flow, it is impossible to understand how information is generated, propagated and dissipated in control systems.

To make progress on these challenging problems, we restrict attention to linear systems as models of plants. Linear systems are the first order approximation of general systems. More importantly, when the systems are linear we can expect to apply the well-developed mathematical tools in linear algebra. For these reasons, in classical control theory linear systems have been used as the first step towards general systems.



(a)



(b)

Figure 1.5: (a) Closed-Loop Control System diagram of “Control over Real Erasure Channels” and (b) Equivalent Open-Loop Estimation System diagram of “Intermittent Kalman Filtering”

1.1 Intermittent Kalman Filtering

In Chapter 2, we will first study intermittent Kalman filtering [95] as a simple toy for abstracting an important piece of control-over-communication-channel problems. As mentioned above, in modern distributed systems, the observer and controller can be located in separate places and connected by unreliable wireless communication channels. Characterizing the tradeoff between control performance and communication channel reliability is a common interest in both control and communication. One of the simplest model for unreliable communication channel is the real erasure channel shown in Figure 1.5a, where the transmitted packet can be lost with a certain probability.

This real erasure channel has both practical and theoretic importance. Practically, the real erasure channel is the simplest toy model for channel fading in wireless communication and packet losses in networks [90]. Theoretically, this system shows that the classical Shannon capacity is not enough to measure information flow for control [87]. Since there is no additive noise and no power constraints on the channel, the classical Shannon capacity of the real erasure channel is infinite. However, the analysis of its performance shows the system can be unstable for sufficiently large erasure probability. Furthermore, [27] found that the maximum erasure probability that systems can tolerate is upper and lower bounded in terms of the largest eigenvalue and the product of all eigenvalues respectively. Therefore, the questions is "Are states amplified through all eigenvalues, or only one or a subset of eigenvalues?" In other words, how do the subspaces of the plant interact with each other when they play a role as information source? Therefore, a proper understanding of this simple model can be a key to understand nature of plants as the sources of control information flows.

The control-estimation separation principle [55] suggests looking at the open-loop estimation system shown in Figure 1.5b as a simplified problem of the original closed-loop control system, focusing on the pure estimation part of the problem. Furthermore, by restricting the observer to be linear time-invariant, the optimal control problem reduces to a seemingly simple variation of Kalman filtering. Formally, the resulting intermittent Kalman filtering problem is written as follows:

$$\mathbf{x}[n+1] = \mathbf{A}\mathbf{x}[n] + \mathbf{w}[n]$$
$$\mathbf{y}[n] = \beta[n](\mathbf{C}\mathbf{x}[n] + \mathbf{v}[n])$$

The vector $\mathbf{x}[n]$ models the states of the plant, and the estimator try to estimate the states based on the observation $\mathbf{y}[n]$. Just as in classical control theory, Gaussian random variables $\mathbf{w}[n]$ and $\mathbf{v}[n]$ are introduced to model uncertainty in the control system. However, unlike classical control problems, Bernoulli random variables $\beta[n]$ are also introduced to model the unreliability in the communication channel.

Even though Kalman filtering gives the optimal estimator, the analysis of its performance is still beyond our understanding. Even the simple but fundamental question "When can we stabilize a plant over an erasure channel?" had been open for years. In Chapter 2, we will answer this question by taking a different approach from the existing literature. While the existing literature attempted to extend Lyapunov stability [95, 27], we generalize observability concepts to definitively answer this question.

We conceptualize the states as the source of control information flow. Again, the key question is to what extent the states of the plant interact as information sources [27]. To answer this question, we introduce a new notion called *eigenvalue cycles*. Borrowing linear algebra concepts, we consider different subspaces of the states as different messages at the source, and measure the size of the messages by the rank of the subspaces. Then, we prove that each subspace of the states can be separated as long as states do not belong to the same eigenvalue cycles, i.e. they do not interact with each other much except when eigenvalue cycles are present. Thus, the original system can be divided into subsystems with smaller dimensions.

The observability gramian of the system can be thought of as a channel which conveys information about the system. Since the amount of source information can be measured by the rank, the rank of the observability gramian has to be large enough to convey enough information about the states. Based on this intuition, we analyze the stopping time until we get enough information about the states, which leads to the characterization of observability of the system.

$$
\begin{aligned}
&x[n+1] = Ax[n] + B_1 u_1[n] + \cdots + C_v u_v[n] \\
&y_1[n] = C_1 x[n] \\
&\vdots \\
&y_v[n] = C_v x[n]
\end{aligned}
$$

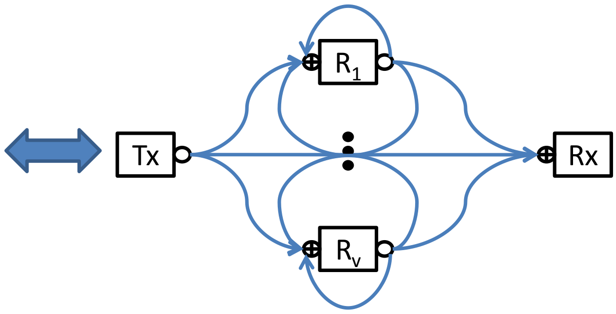


Figure 1.6: Equivalence between Stabilizability of Distributed System with LTI controllers and Capacity of Relay Communication Network

Furthermore, to mathematically justify these insights, we make novel use of information-theoretic ideas. To justify that the original system can be separated into subsystems, we adapt successive decoding [21] and function decoding [74] ideas from modern network information theory. To analyze the stopping time for the observability gramian, we apply large-deviation ideas [24] of information theory.

One of the most counter-intuitive consequences from these insights is that nonuniform sampling can dramatically increase the system robustness. Only the periodicity of the system can make the observability gramian extra-susceptible to becoming rank deficient. However, this periodicity of the system can be easily broken by introducing non-uniform sampling at the observer. With non-uniform sampling, the interaction between the subspaces of the plant can be alleviated. Therefore, the original multi-dimensional system behaves like a collection of simple scalar systems, and the system robustness to channel unreliability can be greatly improved.

This result shares the same spirit as that of compressed sensing [15, 25] where nonuniform sampling or "unstructured" observation matrices are required for optimal recovery. Practically, this idea of nonuniform sampling might be easily implementable. Theoretically, this result also hints at a new general notion of stochastic observability.

1.2 Network Coding meets Decentralized Control

To a centralized observer, the subspaces of the plant do not interact except when eigenvalue cycles make a particular subspace more fragile in terms of the reliability needed. So what about in distributed control systems? Are the different subspaces kept separate? To answer this question, in Chapter 3, we will consider distributed control systems with multiple controllers. We will see that the subspaces associated with different eigenvalues can still be separated, and the amount of information in linear systems can still be measured by the rank of subspaces. As mentioned above, when there are multiple controllers, the information flows between controllers are much more complicated,

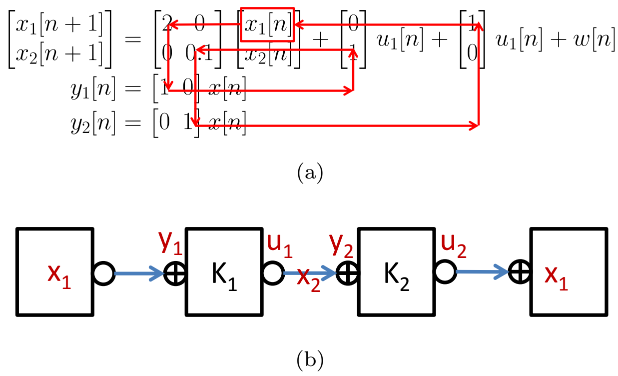


Figure 1.7: (a) An example of distributed control system and (b) The corresponding information flow to stabilize $x_1[n]$

and even finding the flows is not trivial. To simplify the situation, in Chapter 3, we will consider distributed control systems without communication networks. Even without communication networks, the controllers can "implicitly" communicate through the plant. This implicit communication can be much harder to understand than explicit communication via dedicated communication networks.

To understand information flows in distributed systems, we will take a unified approach to control and communication theory. We intuitively believe that when control systems are stabilized, there should be corresponding information flows. As shown in Figure 1.6, we will discover an interesting relationship between distributed linear control systems and linear relay communication networks by considering both as linear systems. We will find an algorithm that relates implicit information flows in control systems with explicit information flows in communication systems.

Recently, the information theory community discovered a new paradigm of *network coding* in understanding information flows in communication networks [1]. In classical communication networks, relays only route their observed signals, while in the network-coding paradigm, relays are allowed not only to route but also to process the signals. Moreover, the information theory community found that there can be a significant gain by allowing such processing [1].

However, this paradigm of processing observations is not a new idea in control theory. All linear controllers inherently mix their observations. Therefore, we can suspect that there might fundamental relationships between control and communication theory.

First, we take a system-theoretic approach to network communication problems. Classically, communication networks are represented by graphs. The flows of graphs are defined as the amount of commodity that we can transfer through the graph, and the cut-values of graphs are defined as the number of edges that go through the cut. Then, the famous graph-theoretic mincut-maxflow theorem [31, 28] reveals that the maximum flow of the graph is equal to the minimum cut of the graph. Furthermore, it was well-known that the mincut is achievable by a static routing scheme, which can be efficiently found by the Ford-Fulkerson algorithm.

However, this graph-theoretic representation of communication networks turns out to be insufficient to model "wireless" communication networks where the signals can be superposed and broadcast by nature. Instead, the connections between the nodes in the network need be modeled by LTI filters. Then, the classical graph-theoretic mincut-maxflow theorem also generalizes to these LTI communication networks [6]. The flow of LTI networks is the rank of the transfer functions, and the cut value of LTI networks is the rank of the channel matrix of the cut. In other words, the information can be measured by the rank of a subspace, which is consonant with the insight of Chapter 2. Then, the maxflow of LTI networks is still equal to the mincut [6]. However, the existing proofs [6, 36, 2, 112] heavily depended on the so-called network unfolding idea [1, 6] which converts general topology networks to layered networks by introducing duplicated nodes over time. As a result, when we fold the layered network back to the original network, even time-invariant schemes in the layered networks become time-varing schemes. It was not clear that we can achieve the mincut of the network even if we restrict the relays to use time-invariant static mix-route schemes.

The difficulty in justifying such a theorem is because the topology of LTI networks can be arbitrarily complicated and can include cycles. To handle this difficulty, we consider LTI communication networks as linear systems, and adapt the state-space representation idea from control theory (linear system theory). We find an algorithm that converts an arbitrary LTI network to a standardized single-hop relay network without changing the time-invariantness of systems. Based on this algorithm, we prove that there exists a static LTI relay scheme that achieves the mincut of LTI networks with static LTI channel operations. In fact, this algorithm can be thought of as a "canonical" state-space representation for LTI networks. Just as we can write all LTI control systems in state-space representation, we can also covert all LTI networks into a standardized single-hop relay network. Furthermore, for general network communication problems like broadcast and unicast, it turns out that unwanted messages at receivers can be modeled as secrecy constraints after the conversion of the networks.

Then, we apply communication theory to understand implicit information flows in distributed control systems. We consider the stabilizability condition for distributed linear systems with LTI controllers. Formally, the system can be written as

$$\begin{aligned}
&\mathbf{x}[n] = \mathbf{A}\mathbf{x}[n] + \mathbf{B_1}\mathbf{u_1}[n] + \cdots + \mathbf{B_v}\mathbf{u_v}[n] \\
&\mathbf{y_1}[n] = \mathbf{C_1}\mathbf{x}[n] \\
&\vdots \\
&\mathbf{y_v}[n] = \mathbf{C_v}\mathbf{x}[n].
\end{aligned}$$

Here, $u_i[n]$ is $(h_i * y_i)[n]$ where $h_i[n]$ is the causal LTI impulse response of the ith controller and $*$ stands for convolution.

It has been well-known that the stabilizability can be characterized using the concept called *fixed modes* [104]. λ is called a fixed mode of system if $\lambda \in \cap_{K_i} \sigma(A + \sum_{1 \leq i \leq v} B_i K_i C_i)$ where $\sigma(\cdot)$

is the set of eigenvalues of the matrix. In [104], it was shown that if there exists a unstable fixed mode, the system is unstabilizable. In [4], an equivalent condition was discovered in a matrix rank form, which does not involve control design parameters K_i in the characterization. It was proved that λ is a fixed mode if and only if $\min_{V \subseteq \{1,\cdots,v\}} \text{rank} \begin{bmatrix} A - \lambda I & B_V \\ C_{V^c} & 0 \end{bmatrix} < \text{rank}(A)$.

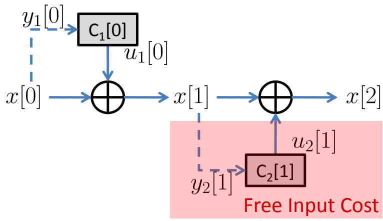
In Chapter 3, we revisit this result and leverage it to reveal the information flow required to stabilize a system. Figure 1.7a shows a descriptive example of implicit communication to stabilize the system. As we can see, the state $x_1[n]$ is only observable by controller 1 while controllable by controller 2. Therefore, to stabilize the state $x_1[n]$, the controller 1 has to communication to the controller 2 through the state $x_2[n]$. Figure 1.7b shows the corresponding information flow to stabilize $x_1[n]$.

The source of the information flow can be thought of as $x_1[n]$. Then, the information is relayed through controller 1 and 2, which are connected by the channel $x_2[n]$. Finally, it arrives at the destination $x_1[n]$. Thus, the source and destination of information flow is the state to stabilize. The controllers are the relays of the network. The remaining states can be thought of as the channel. This answers the basic question we started with: the other states act as information conduits.

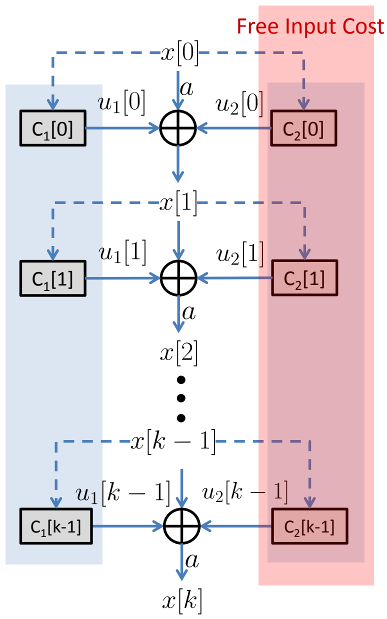
Furthermore, in Chapter 3, we will see that the state is stabilizable if and only if the corresponding communication networks have enough capacity. In other words, for a given unstable eigenvalue, the minimum information required to stabilize that eigenvalue is the rank of the corresponding subspaces (the number of Jordan blocks associated with the eigenvalue). Furthermore, we can construct a relay communication network by considering the remaining states as channels and the controllers as relays. Then, it can be shown that the states corresponding to the unstable eigenvalue are stabilizable if and only if the mincut of the constructed relay network is larger than the rank of the unstable subspaces (the number of Jordan blocks). This result justifies the intuition that to stabilize the control system there has to be a corresponding information flow. More importantly, the controllers "communicate" with each other via robust network coding [52].

1.3 An approximate solution to the decentralized two-controller infinite-horizon scalar LQG problem

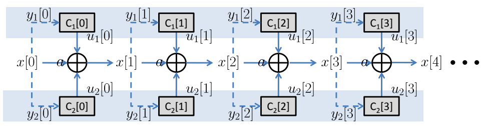
The linear story definitely establishes that information must flow between controllers in distributed LTI systems. The remaining question is "How much do information flows affect the system performance?" To answer this equation, in Chapter 4 and 5, we study optimal control. To study optimal performance, we must relax LTI controller constraints and allow arbitrarily complicated nonlinear or time-varying controllers. By studying optimal control performance, we expect to develop a mathematical and analytic way to quantify control information flows, and understand the impact of "implicit" communication on control performance. This is practically important because



(a)



(b)



(c)

Figure 1.8: (a) Witsenhausen's counterexample, (b) Generalized MIMO Witsenhausen's counterexample, and (c) infinite-horizon LQG control problem.

it can help to understand what are the bottlenecks to performance, and the extent to which these bottlenecks are informational. Then, the use of explicit communication channels might help reach better performance by bypassing those bottlenecks. Since general problems are too difficult, we focus on the simplest nontrivial problem which shows the basic issue —– the infinite-horizon LQG (linear quadratic Gaussian) problem with a scalar plant and two controllers. Formally, the system is given as follows.

$$
\begin{aligned}
&x[n+1] = ax[n] + u_1[n] + u_2[n] + w[n]\\
&y_1[n] = x[n] + v_1[n]\\
&y_2[n] = x[n] + v_2[n]
\end{aligned}
$$

where $w[n]$, $v_1[n]$, $v_2[n]$ are Gaussian random variables. $u_1[n]$, $u_2[n]$ must be causal functions of $y_1[n]$, $y_2[n]$ respectively. The control objective is minimizing the following long-term average cost.

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{1\le n\le N} q\mathbb{E}[x^2[n]] + r_1\mathbb{E}[u_1^2[n]] + r_2\mathbb{E}[u_2^2[n]]$$

The main inspiration is an interesting relationship between Witsenhausen's counterexample [108] and the infinite-horizon LQG problem. Figure 1.8a shows a conceptual diagram of Witsenhausen's counterexample. Witsenhausen's counterexample is also a LQG (linear quadratic Gaussian) problem, but with a (very) finite time-horizon. The first controller acts at the first time step, the second controller acts on the second, and then the system terminates. Thus, the first controller can embed information about its observation in its control input to signal to the second controller [45]. This control-communication dual role of the control input is the crux of the problem, and Witsenhausen's counterexample has been known as the simplest intractable counterexample in distributed control [77, 59].

However, in [39] its relationship with a modern communication problem (dirty paper coding) was revealed. Based on this connection, a nonlinear signaling strategy with theoretical performance guarantees was proposed. By adapting large deviation ideas [24] in information theory, [37] showed that the proposed strategy is approximately optimal.

In Chapter 4 and 5, we will leverage this understanding to infinite-horizon problems. As we can see in Figure 1.8b, the original Witsenhausen's counterexample can be generalized to MIMO Witsenhausen's counterexamples by extending the time-horizon and introducing more observations and control inputs. As we can easily see in Figure 1.8c, these MIMO Witsenhausen's counterexamples are sub-blocks of the infinite-horizon LQG problem. Therefore, we explore this intuition and find a set of constant-ratio-optimal strategies.

To make the connection between MIMO-Witsenhausen and infinite-horizon problems rigorous, we came up with a simple but powerful idea of geometric slicing of problems, which we believe to be the proper way of generalizing information-theoretic cutset bounds to dynamic programming contexts. More importantly, an extensive relationship between wireless communication and decentralized control problems becomes revealed. Conventionally, wireless communication theory divides cases into high-SNR(signal-to-noise ratio) and low-SNR [99]. We discover that the control problems can be also divided into fast-dynamics and slow-dynamics according to the eigenvalues of the system. The implicit information flow in fast(slow) dynamic systems is parallel to the explicit information flow in high(low)-SNR communication systems.

In the fast-dynamics case, we relate the problem to a binary deterministic model [6] first proposed to study information flow in relay communication systems. We conceptualize each bit level of the scalar state as different subspace. Then, two controllers with different observation noise levels can be thought of as observing different subspaces. Thus, the information from the scalar state to controllers can still be measured by the rank of subspaces. The strategy which utilizes the maximum rank of these information turns out to be approximately optimal. Furthermore, in this sense, this control strategy can be thought of as a parallel to the maximum d.o.f. strategy in wireless communication.

In the slow-dynamics case, the maximum-ratio-combining [99] of the observations turns out to be crucial. Therefore, Kalman filtering of the observations is necessary to achieve constant-ratio optimality, while implicit communication between the two controllers is not required. In this sense, this control strategy parallels with the maximum-ratio-combining decoder in wireless communication which exploits the power gain of the signals.

1.4 Future Research

By studying these three simple problems, we begin to understand the nature of information flow for control. Control information flow has its own unique features, but also has similarity to communication information flows. Therefore, current understanding of communication information flows can help to understand control information flows. Especially, we see a striking parallelism between information flows in distributed control and wireless communication systems, and we expect more extensive relationships between control and communication theories will be revealed in future work. Therefore, it will be worth studying how the concepts and ideas from one theory can be properly ported to the other. For example, the information-theoretic-secrecy concept [94] in communication has to be properly converted to secure distributed control [16]. The scaling laws in large wireless communication systems [40] may lead to scaling laws for large distributed control systems. The interference-alignment idea [14, 63] in wireless communication theory is one of the fundamental ideas that has to be infused into distributed control theory. Meanwhile, dynamic programming [11] and delay [85] concepts in control theory also have to be integrated into communication theory. We strongly believe that by studying such relationships we can eventually come up with a unified theory for control and communication. Based on this unified theory, we will truly understand the nature of distributed systems, which will lead to novel and efficient distributed control system designs in practice.

Chapter 2

Intermittent Kalman Filtering

2.1 Introduction

Unlike classical control systems where the controller and the plant are closely located or connected by dedicated wired links, in post-modern systems the controllers and plants can be located far apart and thus control has to happen over communication channels. In other words, there is an observer which can only observe the plant but cannot control it. There is a separate actuator which can only control the plant but cannot observe it. The observer and actuator are connected by a communication channel. Therefore, to control the plant the observer has to send information about its observation to the actuator through the communication channel. Understanding the tradeoff between control performance and communication reliability or finding the optimal controller structures become the fundamental questions to build such post-modern control systems.

Not only practically, but also philosophically, control-over-communication-channel problems are important. When we are controlling systems, there is a corresponding life cycle of information. In other words, the uncertainty or new information is generated and disturbs the plant. This information is propagated to the controller as the controller observes the plant. Finally, when the controller controls the system by removing the uncertainty, the information is dissipated. It is conceptually very important to understand and quantify these information flows which naturally occur as we control systems. In control-over-communication-channel systems, all the information for control has to flow through the communication channel. Therefore, by relating the communication channels with the control performance, we can measure how much information has to flow to achieve a certain control performance.

The theoretical study of control-over-communication-channel problems was pioneered by Baillieul [9] and Tatikonda *et al.* [97]. They restricted the communication channels to noiseless rate-limited channels, and asked what the minimum rate of the channel is to stabilize the plant. They found that the rate of the channel has to be at least the sum of the logarithms of the unstable eigenvalues, and indeed it is sufficient. This fact is known as the data-rate theorem. Later, Nair [70] relaxed the bounded disturbance assumption that they had to Gaussian disturbances, and proved that the same data-rate theorem holds.

However, an important question was whether we can reduce noisy communication channels to noiseless channels with the same Shannon capacity, i.e. whether the classical notion of Shannon capacity is still appropriate when the channel is used for control. In [86], Sahai *et al.* found the answer for this question is no. Intuitively, since the system keeps evolving in time, not only the rate but also the delay of communication is important. Since Shannon capacity ignores the delay issue, it is insufficient for understanding information flows for control. Thus, they proposed a new notion of *anytime capacity* which captures the delay of communication. The stabilizability condition for noisy communication channels with feedback[1] was characterized by anytime capacity.

Since then, researchers have accumulated lots of papers [44, 90, 71, 64, 42, 118, 120] which consider various generalized and related problems. However, still most of the problems are wide open, and the *intermittent Kalman filtering* problem which we will study in this chapter had been one of them. In [95], Sinopoli *et al.* considered 'control over real erasure channels' which can be thought as a special case of [86], but with a structural constraint on controller design.

Figure 2.1 shows the system diagram for control-over-real-erasure-channels. The observer makes observations about the plant, and then uncodedly transmits its observation through the real erasure channel. The real erasure channel drops the transmitted signal with a certain probability but otherwise noiselessly transmits the signal. Finally, based on the received signals from the channel, the controller generates its control inputs to stabilize the system.

The situation that this problem is modeling is that of control over a so-called *packet drop channel*. A memoryless observer samples the output of an unstable continuous-time system, quantizes this sample to a sufficient number of bits, binds the resulting bits into a single packet, and transmits the packet to the controller through a communication system. Due to network congestion or wireless fading, the transmitted packet may be lost[2] with a certain probability and this packet erasure process is further simplified to be i.i.d. The problem is designed to focus attention on the delay/reliability effect of losing packets and so the number of bits per packet (capacity) is unconstrained. The main problem is finding what is the maximum tolerable erasure probability keeping the system stable.

The *linearity* and *memorylessness* of the observer is at the heart of what Sinopoli *et al.* are trying to model. Otherwise, the earlier results of [87] immediately reveal that the critical erasure probability for stabilizability only depends on the magnitude of the largest eigenvalue of the plant. However, to achieve the minimal erasure probability shown in [87], the observer and controller design

[1] By introducing feedback, they reduced the problem to one with nested information structure [109] which is known to be much easier to solve in decentralized control theory. Especially, when the driving disturbance for the plant has bounded support, the plant can be used as implicit noiseless feedback channel. [87]

[2] Such losses need not come from network effects — they could also occur because of sensor occlusion or otherwise at the sampling time itself. That is why the issue of intermittent observations needs to be studied on its own.

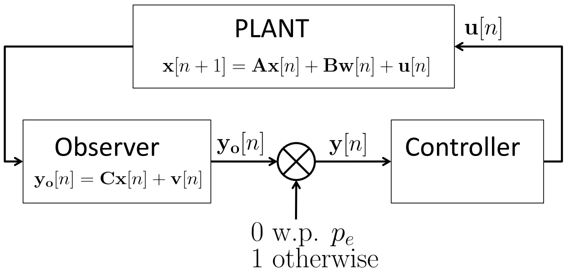


Figure 2.1: Closed-loop system for 'control over real erasure channels'. Here, the observer just passes its observation to the channel without any coding.

has to be quite complicated and may not be realistic in practice. Therefore, it is practically and theoretically important to understand how much the control performance degrades when we impose linear observer and controller constraints.

In this chapter, we will see that the degradation of stabilizability due to linear constraints fundamentally comes only from the periodicity of the system. Nonuniform sampling is proposed as a simple way to force the system to behave aperiodically. Therefore, by using linear controllers in a junction with nonuniform sampling, we can expect a significant performance gain and indeed recover the optimal stabilizability condition over all possible controller designs.

Furthermore, by the estimation-control separation principle [55], the closed-loop control system can be reduced to an equivalent open-loop estimation problem [90]. Figure 2.2 shows the resulting open-loop estimation system so-called *intermittent Kalman filtering* [95]. As before, the sensor uncodedly transmits its observation to the real erasure channel. Then, the estimator tries to estimate the state based on its received signals. We refer to [90] for a literature review and practical applications of the problem.

This chapter is organized as follows: First, we formally state the problem in Section 2.2. Then, we introduce some definitions in Section 2.3. In Section 2.4, we consider intermittent observability as a connection of stability and observability. From this, we distinguish our approach to the previous approaches. In Section 2.5, we introduce some intuition for the characterization of the intermittent observability by using representative examples. In Section 2.6, we formally define eigenvalue cycles and characterize the intermittent observability. In Section 2.7, we discuss how nonuniform sampling can break eigenvalue cycles and significantly improve the robustness of the intermittent Kalman filtering. Finally, Section 2.8 gives the proof of the main results.

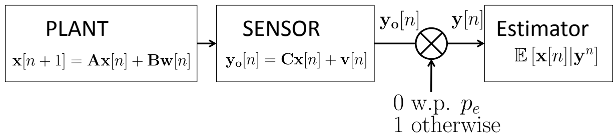


Figure 2.2: System diagram for 'intermittent Kalman filtering'. This open-loop estimation system is equivalent to the closed-loop control system of Figure 2.1. Like Figure 2.1, the sensor bypasses its observation to the channel without any coding.

2.2 Problem Statement

Formally, the intermittent Kalman filtering problem is formulated as follows in discrete time:

$$\mathbf{x}[n+1] = \mathbf{A}\mathbf{x}[n] + \mathbf{B}\mathbf{w}[n] \tag{2.1}$$

$$\mathbf{y}[n] = \beta[n]\left(\mathbf{C}\mathbf{x}[n] + \mathbf{v}[n]\right). \tag{2.2}$$

Here n is the non-negative integer-valued time index and the system variables can take on complex values — i.e. $\mathbf{x}[n] \in \mathbb{C}^m, \mathbf{w}[n] \in \mathbb{C}^g, \mathbf{y}[n] \in \mathbb{C}^l, \mathbf{v}[n] \in \mathbb{C}^l$. $\mathbf{A} \in \mathbb{C}^{m\times m}$, $\mathbf{B} \in \mathbb{C}^{m\times g}$ and $\mathbf{C} \in \mathbb{C}^{l\times m}$. The underlying randomness comes from the initial state $\mathbf{x}[0]$, the persistent driving disturbances $\mathbf{w}[n]$, the observation noises $\mathbf{v}[n]$ and the Bernoulli packet-drops $\beta[n]$. $\beta[n] = 0$ with probability p_e. $\mathbf{x}[0]$, $\mathbf{w}[n]$ and $\mathbf{v}[n]$ are jointly Gaussian.

The objective is to find the best causal estimator $\widehat{\mathbf{x}}[n]$ of $\mathbf{x}[n]$ that minimizes the mean square error (MMSE) $\mathbb{E}[(\mathbf{x}[n] - \widehat{\mathbf{x}}[n])^\dagger(\mathbf{x}[n] - \widehat{\mathbf{x}}[n])]$, i.e. $\widehat{\mathbf{x}}[n] = \mathbb{E}[\mathbf{x}[n]|\mathbf{y}^n]$. We assume that the statistics of all random variables are known to the estimator. If $\mathbf{x}[0]$, $\mathbf{w}[n]$ and $\mathbf{v}[n]$ do not have zero mean, the estimator can properly shift its estimation. Thus, without loss of generality, $\mathbf{x}[0], \mathbf{w}[n]$ and $\mathbf{v}[n]$ are assumed to be zero mean. $\mathbf{x}[0], \mathbf{w}[n]$ and $\mathbf{v}[n]$ are independent and have uniformly bounded second moments so that there exists a positive σ^2 such that

$$\begin{aligned}
&\mathbb{E}[\mathbf{x}[0]\mathbf{x}[0]^\dagger] \preceq \sigma^2\mathbf{I} \\
&\mathbb{E}[\mathbf{w}[n]\mathbf{w}[n]^\dagger] \preceq \sigma^2\mathbf{I} \\
&\mathbb{E}[\mathbf{v}[n]\mathbf{v}[n]^\dagger] \preceq \sigma^2\mathbf{I}.
\end{aligned} \tag{2.3}$$

To prevent degeneracy, we also assume that there exists a positive σ'^2 such that [3]

$$\begin{aligned}
&\mathbb{E}[\mathbf{w}[n]\mathbf{w}[n]^\dagger] \succeq \sigma'^2\mathbf{I} \\
&\mathbb{E}[\mathbf{v}[n]\mathbf{v}[n]^\dagger] \succeq \sigma'^2\mathbf{I}.
\end{aligned} \tag{2.4}$$

[3]The second condition on $\mathbf{v}[n]$ may seem redundant, and $\mathbf{v}[n] = 0$ is enough since at each time the new disturbance $\mathbf{w}[n]$ is added. However, when $\mathbf{v}[n] = 0$, we can make the following counterexample in which the estimation error of the state is bounded even if the system matrices $(\mathbf{A}, \mathbf{C})$ are not observable: $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \mathbf{C} = \begin{bmatrix} 0 & 1 \end{bmatrix}$. Thus, this assumption is usually kept in the analysis of Kalman filtering including [55, p.100].

Under these assumptions we call (2.1) and (2.2) an *intermittent system.*

Definition 2.1. *The linear system equations* (2.1) *and* (2.2) *with the second moment conditions* (2.3) *and* (2.4) *are called an* ***intermittent system*** $(\mathbf{A}, \mathbf{B}, \mathbf{C})$*, or an* ***intermittent system*** $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ ***with erasure probability*** p_e *when we only want to specify the erasure probability, or an* ***intermittent system*** $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \sigma, \sigma')$ ***with erasure probability*** p_e *when we specify the upper and lower bounds on disturbances as well.*

We say that an intermittent system is *intermittent observable* if the MMSE is uniformly bounded for all time.

Definition 2.2. *An intermittent system* $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \sigma, \sigma')$ *with erasure probability* p_e *is called* ***intermittent observable*** *if there exists a casual estimator* $\widehat{\mathbf{x}}[n]$ *of* $\mathbf{x}[n]$ *such that*

$$\sup_{n \in \mathbb{Z}^+} \mathbb{E}[(\mathbf{x}[n] - \widehat{\mathbf{x}}[n])^\dagger (\mathbf{x}[n] - \widehat{\mathbf{x}}[n])] < \infty.$$

Before we discuss truly intermittent cases, let's consider two extreme cases, when $p_e = 1$ and $p_e = 0$, to get some insight into the problem. When $p_e = 1$, the estimator does not have any observations. As a result, the system can be intermittent observable if and only if the system itself is stable. On the other hand, when $p_e = 0$, the estimator has all the observations without any erasures. Intermittent observability reduces to observability. Thus, intermittent observability can be understood as a new concept which interpolates two core concepts of linear system theory: stability and observability.

Moreover, in intermittent systems, we can see the monotonicity of performance with the erasure probability p_e. A process with higher erasure probability can be simulated from a process with lower erasure probability by randomly dropping the observations. Therefore, it is obvious that the average estimation error is an increasing function on p_e. Especially, if we consider an unstable but observable system, when $p_e = 1$ the estimation error goes to infinity, and when $p_e = 0$ the estimation error is bounded. Therefore, between 1 and 0 there must be a threshold on p_e when the estimation error first becomes infinity.

Theorem 2.1 (Theorem 2 of [95])**.** *Given an intermittent system* $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \sigma, \sigma')$ *with erasure probability* p_e*, let* $(\mathbf{A}, \mathbf{B})$ *be controllable*[4]*,* $\sigma < \infty$*, and* $\sigma' > 0$*. Then, there exists a threshold* $p_e^\star$*, such that for* $p_e < p_e^\star$ *the intermittent system* $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \sigma, \sigma')$ *with erasure probability* p_e *is intermittent observable and for* $p_e \geq p_e^\star$ *the intermittent system* $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \sigma, \sigma')$ *with erasure probability* p_e *is not intermittent observable.*

Therefore, the characterization of intermittent observability reduces to the characterization of the critical erasure probability $p_e^\star$. For characterizing the critical erasure probability, we can consider it as a generalization of either stability or observability.

[4]See Definiton 2.3 for controllability.

In [95], Sinopoli *et al.* thought of intermittent observability as a generalization of stability. Based on Lyapunov stability, they could find a lower bound on the critical erasure probability in a LMI (linear matrix inequality) form. However, this bound is not tight in general and does not give any insight into the solution. A more intuitive bound can be found in [27].

Theorem 2.2 (Corollary 8.4 of [27]). *Given an intermittent system* $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \sigma, \sigma')$ *with erasure probability* p_e*, let* $(\mathbf{A}, \mathbf{B})$ *be controllable,* $\sigma < \infty$*,* $\sigma' > 0$*, and* $(\mathbf{A}, \mathbf{C})$ *be observable. Then,*

$$\frac{1}{\prod_i |\lambda_i|^2} \leq p_e^\star \leq \frac{1}{|\lambda_{max}|^2},$$

where λ_i *are the unstable eigenvalues of* $\mathbf{A}$ *and* λ_{max} *is the one with the largest magnitude.*

Therefore, the critical erasure probability characterization boils down to understanding where the gap between $\frac{1}{\prod_i |\lambda_i|^2}$ and $\frac{1}{|\lambda_{max}|^2}$ comes from.

In [113], Mo and Sinopoli found two interesting cases that give further insight into this question. The first is when $\mathbf{A}$ is diagonalizable and all eigenvalues of $\mathbf{A}$ have distinct magnitudes — then the critical erasure probability is $\frac{1}{|\lambda_{max}|^2}$ just it would be in the formulation of [87]. The second case is when $\mathbf{A} = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$ and $\mathbf{C} = \begin{bmatrix} 1 & 1 \end{bmatrix}$ — the critical erasure probability is $\frac{1}{\prod_i |\lambda_i|^2} = \frac{1}{2^4}$. This second case showed that the gap is real and requiring packets to be about a scalar observation can have serious consequences.

To extend these cases and solve the general problem, we will apply insights from observability and introduce the new concept of an *eigenvalue cycle*. As a corollary, we show that in the absence of eigenvalue cycles the critical value becomes $\frac{1}{|\lambda_{max}|^2}$. Furthermore, we show that simply by introducing nonuniform sampling to the sensor, eigenvalue cycles can be broken and the critical erasure probability becomes effectively $\frac{1}{|\lambda_{max}|^2}$.

These results can be surprising if we remember that computing random Lyapunov exponents are difficult problems in general [100]. However, the intermittent Kalman filtering problem turns out to have a special structure which makes the problem tractable. Precisely speaking, as we will see in Section 2.5.3, appropriate subspaces of the vector state can be separated asymptotically. To justify such separation, we use ideas from information theory (for example, decoding functions [74] or successive decoding [21]). Therefore, the whole system can in effect be divided into parallel sub-systems. As we will see in Section 2.5.1, each sub-system can be solved using ideas from large deviation theory [24].

2.3 Definitions and Notations

Before we start the formal discussion of the problem, we first have to introduce mathematical definitions and notations.

We will use controllability and observability notions from linear system theory.

Definition 2.3. *For a $m \times m$ matrix $\mathbf{A}$ and a $m \times p$ matrix $\mathbf{B}$, $(\mathbf{A}, \mathbf{B})$ is called controllable if*

$$\mathcal{C} = \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \cdots & \mathbf{A^{m-1}B} \end{bmatrix}$$

is full rank, or equivalently $\begin{bmatrix} \lambda \mathbf{I} - \mathbf{A} & \mathbf{B} \end{bmatrix}$ is full rank for all $\lambda \in \mathbb{C}$. Moreover, we call an eigenvalue λ of $\mathbf{A}$ uncontrollable if $\begin{bmatrix} \lambda \mathbf{I} - \mathbf{A} & \mathbf{B} \end{bmatrix}$ is rank deficient.

Definition 2.4. *For a $m \times m$ matrix $\mathbf{A}$ and a $l \times m$ matrix $\mathbf{C}$, $(\mathbf{A}, \mathbf{C})$ is called observable if*

$$\mathcal{O} = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA^{m-1}} \end{bmatrix}$$

is full rank, or equivalently $\begin{bmatrix} \lambda \mathbf{I} - \mathbf{A} \\ \mathbf{C} \end{bmatrix}$ is full rank for all $\lambda \in \mathbb{C}$. Moreover, we call an eigenvalue λ of $\mathbf{A}$ unobservable if $\begin{bmatrix} \lambda \mathbf{I} - \mathbf{A} \\ \mathbf{C} \end{bmatrix}$ is rank deficient.

We will use Bernoulli processes and geometric random variables from probability theory.

Definition 2.5. *An one-sided discrete-time random process $a[n]$ $(n \geq 0)$ is called a Bernoulli random process with probability p if $a[n]$ are i.i.d. random variables with the following probability mass function (p.m.f.):*

$$\begin{cases} \mathbb{P}(a[n] = 1) = p \\ \mathbb{P}(a[n] = 0) = 1 - p \end{cases}$$

We also call $a[n]$ as a Bernoulli random variable with erasure probability $1-p$. A two-sided Bernoulli random process is defined in the same way except that n comes from the integers.

Definition 2.6. *A random variable $X \in \mathbb{Z}^+$ is called a geometric random variable with probability p if it has a probability mass function $\mathbb{P}\{X = x\} = p(1-p)^x$ for $x \geq 0$. We also call X as a geometric random variable with erasure probability $1 - p$.*

Then, as it is well known, we have the following relationship between Bernoulli random processes and geometric random variables. Let

$$X := \min\{n \in \mathbb{Z}^+ : a[n] = 1 \text{ where } a[n] \text{ is a Bernoulli random variable with probability } p\}.$$

Then, X is a geometric random variable with probability p.

We will also use the following basic notions about matrices.

Definition 2.7. *Given a matrix $\mathbf{A} \in \mathbb{C}^{m \times m}$, $|\mathbf{A}|_{max}$ is the elementwise max norm of $\mathbf{A}$ i.e. $|\mathbf{A}|_{max} = \max_{1 \leq i,j \leq m} |a_{ij}|$.*

Definition 2.8. *Given a matrix* $\mathbf{A} \in \mathbb{C}^{m\times m}$*,* $\dim \mathbf{A}$ *denotes* m*. Given a column vector* $\mathbf{x_1} \in \mathbb{C}^{m\times 1}$ *and a row vector* $\mathbf{x_2} \in \mathbb{C}^{1\times m}$*,* $\dim \mathbf{x_1}$ *and* $\dim \mathbf{x_2}$ *denote* m*.*

Definition 2.9. *Given* $n_i \times n_i$ *matrices* $\mathbf{A_i}$ *for* $i \in \{1, 2, \cdots, m\}$*,* $diag\{\mathbf{A_1}, \mathbf{A_2}, \cdots, \mathbf{A_m}\}$ *is a* $(\sum_{i=1}^m n_i) \times (\sum_{i=1}^m n_i)$ *matrix in the form of* $\begin{bmatrix} \mathbf{A_1} & 0 & \cdots & 0 \\ 0 & \mathbf{A_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{A_m} \end{bmatrix}$.

We also define modulo operation on numbers.

Definition 2.10. *A sequence,* $a_1, a_2, \cdots, a_n$*, is called congruent mod* p *if* $a_i \equiv a_j (mod\ p)$ *for all* i, j*.*

Definition 2.11. *A sequence,* $a_1, a_2, \cdots, a_n$*, is called pairwise incongruent mod* p *if* $a_i \not\equiv a_j (mod\ p)$ *for all* $i \neq j$*.*

Since we will only focus on the scaling behavior, we will use the following definition paralleling big O and big Ω notations in complexity theory.

Definition 2.12. *Consider two real functions* $a(t)$ *and* $b(t)$ *whose common domain is* $T \subseteq \mathbb{R}$*. We say* $a(t) \lesssim b(t)$ *for* t *on* T *if there exists a positive* c *such that* $a(t) \leq cb(t)$ *for all* $t \in T$*.*

We omit the argument and the domain of the above definition, when they are obvious from the context and do not cause confusion.

We will also use an abbreviated notation for a sequence of random variables.

Definition 2.13. *Given a discrete time random variable* $a[0], \cdots, a[n]$*, we denote* $a[n_1], \cdots, a[n_2]$ *as* $a_{n_1}^{n_2}$*, and* $a[0], \cdots, a[n]$ *as* a^n*. Likewise given a continuous time random variable* $b(t)$*, we define* $\mathbf{b}(t_1 : t_2)$ *to be* $\mathbf{b}(t)$ *for* $t_1 \leq t \leq t_2$*.*

2.4 Intermittent Observability as an Extension of Stability

As we mentioned before, the characterization of the critical erasure probability can be considered from two different directions — an extension of stability or an extension of observability. In [95], Sinopoli *et al.* took the first approach, and attempted to characterize the critical erasure probability by the Lyapunov stability condition. Let's first review a property of Schur complements and Lyapunov stability theorem.

Lemma 2.1 (Schur complements)**.** *Let* $\mathbf{X} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^\dagger & \mathbf{C} \end{bmatrix}$ *be a symmetric matrix and* $\mathbf{C}$ *be invertible. Then,* $\mathbf{X} \succ 0$ *if and only if* $\mathbf{C} \succ 0$ *and* $\mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^\dagger \succ 0$*.*

Proof. See [12, p. 650]. □

Theorem 2.3 (Lyapunov Stability Theorem)**.** *Given a linear system* (2.1)*, the following three conditions are equivalent.*
(i) The system is stable.
(ii) $\exists \mathbf{M}, \mathbf{N} \succ \mathbf{0}$ *such that*

$$\mathbf{M} - \mathbf{A}\mathbf{M}\mathbf{A}^{\dagger} = \mathbf{N}.$$

(iii) $\exists \mathbf{M} \succ \mathbf{0}$ *such that*

$$\begin{bmatrix} \mathbf{M} & \mathbf{A}\mathbf{M} \\ \mathbf{M}\mathbf{A}^{\dagger} & \mathbf{M} \end{bmatrix} \succ \mathbf{0}.$$

Proof. The equivalence between (i) and (ii) can be easily found in linear system theory books including [55, p.30] and [17, Theorem 5.D5]. The equivalence between (ii) and (iii) comes from Schur complements in Lemma 2.1 by simply choosing $\mathbf{A} = \mathbf{M}$, $\mathbf{B} = \mathbf{A}\mathbf{M}$ and $\mathbf{C} = \mathbf{M}$. □

Before we consider intermittent observability, let's first characterize the standard observability condition using Lyapunov stability. The fundamental theorem of observability tells that if $(\mathbf{A}, \mathbf{C})$ is observable, the eigenvalues of the closed loop system $\mathbf{A} + \mathbf{K}\mathbf{C}$ can be placed anywhere by a proper selection of $\mathbf{K}$. Based on this, we can characterize observability in terms of Lyapunov stability.

Theorem 2.4. *Given a linear system* (2.1) *and* (2.2) *with* $p_e = 0$*, the following four conditions are equivalent.*
(i) All the unstable modes of $\mathbf{A}$ *are observable.*
(ii) $\exists \mathbf{K}$ *such that* $\mathbf{A} + \mathbf{K}\mathbf{C}$ *is stable.*
(iii) $\exists \mathbf{K}$ *and* $\mathbf{M}, \mathbf{N} \succ \mathbf{0}$ *such that*

$$\mathbf{M} - (\mathbf{A} + \mathbf{K}\mathbf{C})\mathbf{M}(\mathbf{A} + \mathbf{K}\mathbf{C})^{\dagger} = \mathbf{N}.$$

(iv) $\exists \mathbf{K}$ *and* $\mathbf{M} \succ \mathbf{0}$ *such that*

$$\begin{bmatrix} \mathbf{M} & (\mathbf{A} + \mathbf{K}\mathbf{C})\mathbf{M} \\ \mathbf{M}(\mathbf{A} + \mathbf{K}\mathbf{C})^{\dagger} & \mathbf{M} \end{bmatrix} \succ 0.$$

Proof. The equivalence of (i) and (ii) is the fundamental theorem of observability [17, Theorem 8.M3]. The equivalence of (ii), (iii) and (iv) follows from Theorem 2.3. □

Unfortunately, this observability characterization based on Lyapunov stability cannot be generalized for intermittent observability. The main reason is that in intermittent Kalman filtering the optimal estimator does not converge to a linear time-invariant one. In conventional Kalman filtering for linear time-invariant systems, it is well-known that the optimal Kalman filter converges to the linear time-invariant estimator which is known as the *Wiener filter* [107]. In fact, we can directly plug in the Wiener filter gain for the matrix $\mathbf{K}$ of Theorem 2.4. However, when observations are erased, the optimal estimator also depends on the erasure pattern and since the erasure pattern is random and time-varying, the whole system becomes random and time-varying. Therefore, the optimal estimator is also time-varying and does not converge.

In [95], Sinopoli *et al.* wrote the optimal time-varying linear estimator in a recursive equation form. The strictly causal estimator $\widehat{\mathbf{x}}[n] = \mathbb{E}[\mathbf{x}[n]|\mathbf{y}^{n-1}]$, is given as follows:

$$\widehat{\mathbf{x}}[n+1] = \mathbf{A}\widehat{\mathbf{x}}[n] - \mathbf{K_n}(\mathbf{y}[n] - \mathbf{C}\widehat{\mathbf{x}}[n]) \tag{2.5}$$

Here, $\mathbf{K_n}$ depends not only on n but also the history of the $\beta[n]$, and does not converge to a constant matrix in probability. Therefore, in the intermittent Kalman filtering problem it is not possible to find a stability-optimal time-invariant gain $\mathbf{K}$ in Theorem 2.4.

However, we can still force the estimator to be linear time-invariant, and thereby find a sufficient condition for intermittent observability using Lyapunov stability ideas. This is the idea that Sinopoli *et al.* used to find a lower bound on the critical erasure probability in [95]. By restricting the filtering gain to be a linear time-invariant matrix $\mathbf{K}$, we get the following sub-optimal estimator which looks similar to (2.5).

$$\widehat{\mathbf{x}}[n+1] = \mathbf{A}\widehat{\mathbf{x}}[n] - \beta[n]\mathbf{K}(\mathbf{y}[n] - \mathbf{C}\widehat{\mathbf{x}}[n]) \tag{2.6}$$

with $\widehat{\mathbf{x}}[0] = \mathbf{0}$. By analyzing this sub-optimal estimator, Sinopoli *et al.* found the following sufficient condition for intermittent observability. Here, we further prove that their condition is both necessary and sufficient for the sub-optimal estimators of (2.6) to have an expected estimation error uniformly bounded over time.[5]

Theorem 2.5 (Extension of Theorem 5 of [95])**.** *Given an intermittent system* $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \sigma, \sigma')$ *with erasure probability* p_e*, let* $(\mathbf{A}, \mathbf{B})$ *be controllable,* $\sigma < \infty$*, and* $\sigma' > 0$*. Then, the following three conditions are equivalent.*
(i) The system is intermittently observable by the suboptimal estimator of (2.6) *with some* $\mathbf{K}$*.*
(ii) $\exists \mathbf{K}$ *and* $\mathbf{M}, \mathbf{N} \succ \mathbf{0}$ *such that*

$$\mathbf{M} - p_e\mathbf{A}\mathbf{M}\mathbf{A}^\dagger - (1-p_e)(\mathbf{A}+\mathbf{K}\mathbf{C})\mathbf{M}(\mathbf{A}+\mathbf{K}\mathbf{C})^\dagger = \mathbf{N}.$$

[5] This fact is implicitly shown in Elia's paper [27].

(iii) $\exists\mathbf{K}$ *and* $\mathbf{M} \succ \mathbf{0}$ *such that*

$$\begin{bmatrix} \mathbf{M} & \sqrt{1-p_e}(\mathbf{MA}+\mathbf{KC}) & \sqrt{p_e}\mathbf{MA} \\ \sqrt{1-p_e}(\mathbf{MA}+\mathbf{KC})^\dagger & \mathbf{M} & 0 \\ \sqrt{p_e}(\mathbf{MA})^\dagger & 0 & \mathbf{M} \end{bmatrix} \succ \mathbf{0}.$$

Proof. By (2.1), (2.2) and (2.6), we can see that the estimation error follows the following dynamics:

$$\begin{aligned} \mathbf{x}[n+1] - \widehat{\mathbf{x}}[n+1] &= \mathbf{A}\mathbf{x}[n] + \mathbf{B}\mathbf{w}[n] - (\mathbf{A}\widehat{\mathbf{x}}[n] - \beta[n]\mathbf{K}(\mathbf{y}[n] - \mathbf{C}\widehat{\mathbf{x}}[n])) \\ &= \mathbf{A}\mathbf{x}[n] + \mathbf{B}\mathbf{w}[n] - (\mathbf{A}\widehat{\mathbf{x}}[n] - \beta[n]\mathbf{K}(\mathbf{C}\mathbf{x}[n] + \mathbf{v}[n] - \mathbf{C}\widehat{\mathbf{x}}[n])) \\ &= (\mathbf{A} + \beta[n]\mathbf{KC})(\mathbf{x}[n] - \widehat{\mathbf{x}}[n]) + \mathbf{B}\mathbf{w}[n] + \beta[n]\mathbf{K}\mathbf{v}[n]. \end{aligned} \tag{2.7}$$

Denote $(\mathbf{x}[n] - \widehat{\mathbf{x}}[n])$ as $(\mathbf{e}[n]$ and $\mathbf{B}\mathbf{w}[n] + \beta[n]\mathbf{K}\mathbf{v}[n])$ as $\mathbf{w}'[n]$. Then, $\mathbf{w}'[n]$ also has a uniformly bounded variance over time, and (2.7) can be written as

$$\mathbf{e}[n+1] = (\mathbf{A} + \beta[n]\mathbf{KC})\mathbf{e}[n] + \mathbf{w}'[n].$$

Since $\mathbf{e}[n]$ is independent from $\mathbf{w}'[n], \beta[n]$ by causality, the covariance matrix of $\mathbf{e}[n]$ follows the following dynamics:

$$\begin{aligned} \mathbb{E}[\mathbf{e}[0]\mathbf{e}^\dagger[0]] &= \mathbb{E}[\mathbf{x}[0]\mathbf{x}^\dagger[0]], \\ \mathbb{E}[\mathbf{e}[n+1]\mathbf{e}^\dagger[n+1]] &= \mathbb{E}[(\mathbf{A} + \beta[n]\mathbf{KC})\mathbf{e}[n]\mathbf{e}^\dagger[n](\mathbf{A} + \beta[n]\mathbf{KC})^\dagger] + \mathbb{E}[\mathbf{w}'[n]\mathbf{w}'^\dagger[n]] \\ &= p_e\mathbf{A}\mathbb{E}[\mathbf{e}[n]\mathbf{e}^\dagger[n]]\mathbf{A}^\dagger + (1-p_e)(\mathbf{A}+\mathbf{KC})\mathbb{E}[\mathbf{e}[n]\mathbf{e}^\dagger[n]](\mathbf{A}+\mathbf{KC})^\dagger + \mathbb{E}[\mathbf{w}'[n]\mathbf{w}'^\dagger[n]]. \end{aligned} \tag{2.8}$$

Now, we will prove the theorem in three steps.

(1) Condition (i) implies condition (ii).

First of all, by linearity we can prove that the estimation error $\mathbb{E}[\mathbf{e}[n]\mathbf{e}^\dagger[n]]$ is an increasing function of the variance of the underlying random variables.

Thus, if the system is intermittently observable by $\mathbf{K}$, the same system with $\mathbf{x}[0] = 0$, $\mathbf{v}[n] = 0$, $\mathbb{E}[\mathbf{w}[n]\mathbf{w}^\dagger[n]] = \sigma'^2\mathbf{I}$ is also intermittently observable. So set $\mathbf{x}[0] = 0$, $\mathbf{v}[n] = 0$, $\mathbb{E}[\mathbf{w}[n]\mathbf{w}^\dagger[n]] = \sigma'^2\mathbf{I}$ without loss of generality. With these parameters, we have $\mathbb{E}[\mathbf{e}[0]\mathbf{e}^\dagger[0]] = 0$ and $\mathbb{E}[\mathbf{w}'[n]\mathbf{w}'^\dagger[n]] = \sigma'^2\mathbf{B}\mathbf{B}^\dagger$. By the recursive equation in (2.8), we can show that for $n \geq 1$, the covariance matrix of $\mathbf{e}[n]$ can be written as

$$\mathbb{E}[\mathbf{e}[n]\mathbf{e}^\dagger[n]] = \sigma'^2\mathbf{B}\mathbf{B}^\dagger + \sum_{k=1}^{n} \sum_{l \in \{-1,1\}^k} \mathbf{\Delta}_l \mathbf{\Delta}_l^\dagger.$$

where

$$\mathbf{\Delta}_l := (\sqrt{p_e}\mathbf{A})^{\frac{1+l_1}{2}} (\sqrt{1-p_e}(\mathbf{A}+\mathbf{KC}))^{\frac{1-l_1}{2}} \cdots (\sqrt{p_e}\mathbf{A})^{\frac{1+l_k}{2}} (\sqrt{1-p_e}(\mathbf{A}+\mathbf{KC}))^{\frac{1-l_k}{2}} \sigma'\mathbf{B}.$$

Here, $l_i = 1$ means the ith observation was erased and $l_i = -1$ means that the ith observation was not erased.

Here, we can notice that $\mathbb{E}[\mathbf{e}[n]\mathbf{e}^\dagger[n]]$ are positive semidefinite matrices and increasing in n. Furthermore, since the system is intermittently observable by condtion (i), $\mathbb{E}[\mathbf{e}[n]\mathbf{e}^\dagger[n]]$ has to be uniformly bounded over time. Therefore,

$$\bar{\mathbf{M}} := \lim_{n\to\infty} \mathbb{E}[\mathbf{e}[n]\mathbf{e}^\dagger[n]] = \sigma'^2\mathbf{B}\mathbf{B}^\dagger + \sum_{k=1}^{\infty}\sum_{l\in\{-1,1\}^k} \mathbf{\Delta}_l\mathbf{\Delta}_l^\dagger \tag{2.9}$$

must exist even though it involves an infinite sum. Let's define $\mathbf{M}$ and $\mathbf{N}$ as follows:

$$\mathbf{M} := \sigma'^2\mathbf{B}\mathbf{B}^\dagger + \sum_{k=1}^{m-1}\sum_{l=\{-1,1\}^k}(k+1)\mathbf{\Delta}_l\mathbf{\Delta}_l^\dagger + \sum_{k'=m}^{\infty}\sum_{l'=\{-1,1\}^{k'}} m\mathbf{\Delta}_l\mathbf{\Delta}_l^\dagger \tag{2.10}$$

$$\mathbf{N} := \sigma'^2\mathbf{B}\mathbf{B}^\dagger + \sum_{k=1}^{m-1}\sum_{l\in\{-1,1\}^k} \mathbf{\Delta}_l\mathbf{\Delta}_l^\dagger \tag{2.11}$$

where m is the dimension of $\mathbf{A}$ as we defined in Section 2.2. By the definitions of $\bar{\mathbf{M}}$ and $\mathbf{M}$, we can easily see that $m\bar{\mathbf{M}} \succeq \mathbf{M}$. Therefore, $\mathbf{M}$ also exists even though it involves an infinite sum. Furthermore, by the definitions of $\mathbf{M}$ and $\mathbf{N}$, we can easily see that

$$\mathbf{M} \succeq \sigma'^2(\mathbf{B}\mathbf{B}^\dagger + p_e\mathbf{A}\mathbf{B}\mathbf{B}^\dagger\mathbf{A}^\dagger + \cdots + p_e^m\mathbf{A}^m\mathbf{B}\mathbf{B}^\dagger\mathbf{A}^{\dagger m})$$
$$\mathbf{N} \succeq \sigma'^2(\mathbf{B}\mathbf{B}^\dagger + p_e\mathbf{A}\mathbf{B}\mathbf{B}^\dagger\mathbf{A}^\dagger + \cdots + p_e^m\mathbf{A}^m\mathbf{B}\mathbf{B}^\dagger\mathbf{A}^{\dagger m})$$

since the terms in L.H.S. are just subsets of the terms in $\mathbf{M}$ and $\mathbf{N}$.

Thus, we can see that $\mathbf{M} \succ 0$, $\mathbf{N} \succ 0$ since $\begin{bmatrix}\mathbf{B} & \mathbf{A}\mathbf{B} & \cdots & \mathbf{A}^{m-1}\mathbf{B}\end{bmatrix}$ is full rank by the controllability of $(\mathbf{A}, \mathbf{B})$ and all terms $\mathbf{B}\mathbf{B}^\dagger, \cdots, p_e^m\mathbf{A}^m\mathbf{B}\mathbf{B}^\dagger\mathbf{A}^{\dagger m}$ are positive semidefinite. Finally, by the definitions and simple matrix algebra, we can verify that $\mathbf{M}$ and $\mathbf{N}$ satisfy the following relationship:

$$\mathbf{M} = p_e\mathbf{A}\mathbf{M}\mathbf{A}^\dagger + (1-p_e)(\mathbf{A}+\mathbf{K}\mathbf{C})\mathbf{M}(\mathbf{A}+\mathbf{K}\mathbf{C})^\dagger + \mathbf{N}. \tag{2.12}$$

Therefore, $\mathbf{M}$ and $\mathbf{N}$ satisfy condition (ii).[6]

(2) Condition (ii) implies condition (i).

Since $\mathbf{M}$ and $\mathbf{N}$ of condition (ii) are positive definite, we can find a such that $a^2\mathbf{M} \succ \mathbb{E}[\mathbf{x}[0]\mathbf{x}^\dagger[0]]$ and $a^2\mathbf{N} \succ \mathbb{E}[\mathbf{w}'[n]\mathbf{w}'^\dagger[n]]$ for all $n \in \mathbb{Z}^+$. And we can easily see that even if we replace $\mathbf{K}$, $\mathbf{M}$, $\mathbf{N}$ with $\mathbf{K}$, $a^2\mathbf{M}$, $a^2\mathbf{N}$, condition (ii) still holds.

[6] Consider a fixed point equation, $f(x) = xf(x)+g(x)$. There exist multiple $f(x)$ and $g(x)$ that satisfy this equation. For example, $(f(x), g(x)) = (1 + x + x^2 + \cdots, 1)$, $(f(x), g(x)) = (1 + 2x + 2x^2 + \cdots, 1 + x)$, $\cdots$, $(f(x), g(x)) = (1 + 2x + \cdots + (k-1)x^{k-1} + kx^k + kx^{k+1}\cdots, 1 + x + \cdots + x^k)$ all satisfy the equation. Likewise, there are multiple matrices that satisfy the fixed point equation of (2.12). For example, we can easily check that $\bar{\mathbf{M}}$ of (2.9) and $\bar{\mathbf{N}} := \sigma'^2\mathbf{B}\mathbf{B}^\dagger$ satisfy (2.12), i.e. $\bar{\mathbf{M}} = p_e\mathbf{A}\bar{\mathbf{M}}\mathbf{A}^\dagger + (1-p_e)(\mathbf{A}+\mathbf{K}\mathbf{C})(\mathbf{A}+\mathbf{K}\mathbf{C})^\dagger + \bar{\mathbf{N}}$. However, unlike $\mathbf{N}$, $\bar{\mathbf{N}}$ does not have to be positive definite. Thus, the choice of $\bar{\mathbf{M}}$, $\bar{\mathbf{N}}$ is not enough to prove the theorem. Here, we choose $\mathbf{M}$, $\mathbf{N}$ as shown in (2.10), (2.11) as another solution for (2.12). In fact, the choice of coefficient in $\mathbf{M}, \mathbf{N}$ was inspired by the solutions of $f(x) = xf(x) + g(x)$ shown above.

We will prove that $a^2\mathbf{M} \succ \mathbb{E}[\mathbf{e}[n]\mathbf{e}^\dagger[n]]$ for all $n \in \mathbb{Z}^+$ by induction. Since $a^2\mathbf{M} \succ \mathbb{E}[\mathbf{x}[0]\mathbf{x}^\dagger[0]] = \mathbb{E}[\mathbf{e}[0]\mathbf{e}^\dagger[0]]$, the claim is true for $n = 0$. Assume the claim is true for n. Then, from the definition of a and (2.8),

$$\mathbb{E}[\mathbf{e}[n+1]\mathbf{e}^\dagger[n+1]] \prec p_e\mathbf{A}(a^2\mathbf{M})\mathbf{A}^\dagger + (1-p_e)(\mathbf{A}+\mathbf{KC})(a^2\mathbf{M})(\mathbf{A}+\mathbf{KC})^\dagger + a^2\mathbf{N} = a^2\mathbf{M}$$

where the last equality comes from condition (ii). Therefore, the estimation error is uniformly upper bounded by $a^2\mathbf{M}$ when we use the $\mathbf{K}$ of condition (ii) as a gain matrix, and so condition (ii) implies condition (i).

(3) Condition (ii) is equivalent to condition (iii).

Since $\mathbf{M}^{-1} \succ \mathbf{0}$, by Schur complements in Lemma 2.1, condition (ii) is equivalent to

$$\begin{bmatrix} \mathbf{M} - p_e\mathbf{AMA}^\dagger & \sqrt{1-p_e}(\mathbf{A}+\mathbf{KC}) \\ \sqrt{1-p_e}(\mathbf{A}+\mathbf{KC})^\dagger & \mathbf{M}^{-1} \end{bmatrix} \succ \mathbf{0}.$$

Since

$$\begin{aligned}&\begin{bmatrix} \mathbf{M} - p_e\mathbf{AMA}^\dagger & \sqrt{1-p_e}(\mathbf{A}+\mathbf{KC}) \\ \sqrt{1-p_e}(\mathbf{A}+\mathbf{KC})^\dagger & \mathbf{M}^{-1} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{M} & \sqrt{1-p_e}(\mathbf{A}+\mathbf{KC}) \\ \sqrt{1-p_e}(\mathbf{A}+\mathbf{KC})^\dagger & \mathbf{M}^{-1} \end{bmatrix} - \begin{bmatrix} \sqrt{p_e}\mathbf{A} \\ 0 \end{bmatrix} \mathbf{M} \begin{bmatrix} \sqrt{p_e}\mathbf{A}^\dagger & 0 \end{bmatrix}\end{aligned}$$

and $\mathbf{M}^{-1} \succ \mathbf{0}$, we can apply Schur complement again. Thus, condition (ii) is equivalent to

$$\begin{bmatrix} \mathbf{M} & \sqrt{1-p_e}(\mathbf{A}+\mathbf{KC}) & \sqrt{p_e}\mathbf{A} \\ \sqrt{1-p_e}(\mathbf{A}+\mathbf{KC})^\dagger & \mathbf{M}^{-1} & 0 \\ \sqrt{p_e}\mathbf{A}^\dagger & 0 & \mathbf{M}^{-1} \end{bmatrix} \succ \mathbf{0}.$$

Since $\mathbf{M}^{-1} \succ 0$, this condition is again equivalent to

$$\begin{aligned}&\begin{bmatrix} \mathbf{M}^{-1} & 0 & 0 \\ 0 & \mathbf{I} & 0 \\ 0 & 0 & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{M} & \sqrt{1-p_e}(\mathbf{A}+\mathbf{KC}) & \sqrt{p_e}\mathbf{A} \\ \sqrt{1-p_e}(\mathbf{A}+\mathbf{KC})^\dagger & \mathbf{M}^{-1} & 0 \\ \sqrt{p_e}\mathbf{A}^\dagger & 0 & \mathbf{M}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{M}^{-1} & 0 & 0 \\ 0 & \mathbf{I} & 0 \\ 0 & 0 & \mathbf{I} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{M}^{-1} & \sqrt{1-p_e}(\mathbf{M}^{-1}\mathbf{A}+\mathbf{M}^{-1}\mathbf{KC}) & \sqrt{p_e}\mathbf{M}^{-1}\mathbf{A} \\ \sqrt{1-p_e}(\mathbf{M}^{-1}\mathbf{A}+\mathbf{M}^{-1}\mathbf{KC})^\dagger & \mathbf{M}^{-1} & 0 \\ \sqrt{p_e}(\mathbf{M}^{-1}\mathbf{A})^\dagger & 0 & \mathbf{M}^{-1} \end{bmatrix} \succ \mathbf{0}.\end{aligned}$$

Since $\mathbf{M}^{-1} \succ \mathbf{0}$ and $\mathbf{K}$ is an arbitrary matrix, by replacing $\mathbf{M}^{-1}$ by $\mathbf{M}$ and $\mathbf{M}^{-1}\mathbf{K}$ by $\mathbf{K}$ we get condition (iii). □

As we can expect, the conditions of this theorem reduce to those of stability and those of observability when $p_e = 1$ and $p_e = 0$ respectively. One can easily observe that condition (ii) of Theorem 2.5 reduces to condition (ii) of Theorem 2.3 when $p_e = 1$ and condition (iii) of Theorem 2.4

when $p_e = 0$. Likewise, condition (iii) of Theorem 2.5 reduces to condition (iii) of Theorem 2.3 and condition (iv) of Theorem 2.4 respectively.

Even though condition (ii) and (iii) of Theorem 2.5 are equivalent, condition (iii) is preferred since it is given in a LMI (linear matrix inequality) form and convex optimization techniques [12] are applicable. In fact, in [95] Sinopoli *et al.* related condition (iii) with quasi-convex problems.

Since we imposed an additional linear time-invariant constraint on the estimator, Theorem 2.5 gives a lower bound on $p_e^\star$. However, we can conclude that this lower bound is loose in general.[7] Moreover, even for stability, the characterization that the magnitudes of all eigenvalues are less than 1 is much more intuitive than the LMI condition based on Lyapunov stability. Therefore, researchers including [27] and [113] were looking for a tight and intuitive characterization of the critical erasure probability.

2.5 Intermittent observability as an extension of observability: Main Intuition

To reach this goal, we borrow insights from a characterization of observability. $(\mathbf{A}, \mathbf{C})$ is observable if and only if for all $s \in \mathbb{C}$

$$\begin{bmatrix} s\mathbf{I} - \mathbf{A} \\ \mathbf{C} \end{bmatrix} \text{ is full rank.}$$

Moreover, by a similarity transform [17] we can assume that $\mathbf{A}$ is in Jordan form[8] without loss of generality. With this additional assumption, the observability condition can be further simplified.

Theorem 2.6 ([17]). *Consider a linear system with system matrices $(\mathbf{A}, \mathbf{C})$ where $\mathbf{A}$ is given in a Jordan form. For an eigenvalue λ of $\mathbf{A}$, denote $\mathbf{C}_\lambda$ as a matrix whose columns consist of the columns of $\mathbf{C}$ which correspond to the first elements of the Jordan blocks in $\mathbf{A}$ associated with λ. Then, the states associated with λ are observable if and only if the rank of $\mathbf{C}_\lambda$ is equal to the number of Jordan blocks associated with λ. The whole system is observable if and only if all states associated with all eigenvalues are observable.*

---

[7] Numerical computation of the lower bound of Theorem 2.5 is shown in Figure 4 of [95]. For a system with $\mathbf{A} = \begin{bmatrix} 1.25 & 0 \\ 1 & 1.1 \end{bmatrix}$ and $\mathbf{C} = \begin{bmatrix} 1 & 1 \end{bmatrix}$. The numerical simulation shows the lower bound is approximately $\frac{1}{(1.25 \times 1.1)^2} = 0.528\cdots$, while the exact characterization of Theorem 2.7 tells the critical erasure probability is $\frac{1}{1.25^2} = 0.64$.

[8] Throughout the chapter, we will use the Jordan form that induces an upper triangular matrix.

For example, let

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$
$$\mathbf{C} = \begin{bmatrix} \mathbf{c_1} & \mathbf{c_2} & \mathbf{c_3} & \mathbf{c_4} \end{bmatrix}.$$

Then, $\mathbf{C}_2 = \begin{bmatrix} \mathbf{c_1} & \mathbf{c_3} \end{bmatrix}$ and $\mathbf{C}_3 = \begin{bmatrix} \mathbf{c_4} \end{bmatrix}$. The eigenvalue 2 is observable if and only if $\mathbf{C}_2$ is full rank, and the eigenvalue 3 is observable if and only if $\mathbf{C}_3$ is full rank. The whole system with $(\mathbf{A}, \mathbf{C})$ is observable if and only if both eigenvalues are observable.

This characterization reminds us of a *divide-and-conquer* approach. First, divide the observability problem into smaller problems according to the eigenvalues. Then, check whether the smaller sub-problem for each eigenvalue is observable. Finally, the whole system is observable if and only if all the sub-problems are observable.

This suggests applying a divide-and-conquer approach for the characterization of intermittent observability. However, before we apply a divide-and-conquer approach, we first have to answer the following three questions:

(a) What are the minimal irreducible sub-problems?

(b) How can we solve each sub-problem?

(c) How can we combine the answers of the sub-problems?

We will make an exact characterization of intermittent observability by resolving these questions. The concept of eigenvalue cycles appears naturally as the answer of question (a).

Before we answer these questions, let's first start from the simplest case, scalar plants. For simplicity, we will only give hand-waving arguments in this section, and the rigorous justification will be shown in later sections. The basic idea for the characterization of intermittent observability is to consider the dynamics reverse in time. For example, consider the following scalar system: for $n \in \mathbb{Z}^+$,

$$\begin{cases} x[n+1] = 2x[n] + w[n] \\ y[n] = \beta[n]x[n] \end{cases}. \tag{2.13}$$

Here, $x[0] = 0$, $w[n]$ are i.i.d. zero-mean unit-variance Gaussian, and $\beta[n]$ is an independent Bernoulli process with probability $1 - p_e$. Then, we will show that the critical erasure probability $p_e^\star = \frac{1}{2^2}$.

First, we extend the one-sided random process (2.13) to a two-sided process. Let $w[n] = 0$ for $n \in \mathbb{Z}^{--}$ where $\mathbb{Z}^{--}$ implies negative integers, and $\beta[n]$ be a two-sided Bernoulli process with probability $1 - p_e$. Then, we can see that the new two-sided process is equivalent to the original process except that $x[n] = 0, y[n] = 0$ for $n \in \mathbb{Z}^{--}$.

Let $n - S$ be the most recent non-erased observation at time n, i.e. $S := \min\{k \geq 0 : \beta[n-k] = 1\}$. Since $\beta[n]$ is a two-sided Bernoulli process, the stopping time S is a geometric random variable, i.e. $\mathbb{P}\{S = s\} = (1 - p_e){p_e}^s$.

(1) Sufficiency: We first prove that $p_e < \frac{1}{2^2}$ is sufficient for intermittent observability of the example. For this, we analyze the performance of a suboptimal estimator $\widehat{x}[n] = 2^S y[n - S] = 2^S x[n - S]$. Then, the estimation error is upper bounded by

$$\begin{aligned}
\mathbb{E}[(x[n] - \widehat{x}[n])^2] &= \mathbb{E}[\mathbb{E}[(x[n] - \widehat{x}[n])^2|S]] \\
&= \mathbb{E}[\mathbb{E}[(2^S x[n - S] + 2^{S-1} w[n - S] + \cdots + w[n - 1] - 2^S x[n - S])^2|S]] \\
&\leq \mathbb{E}[2^{2(S-1)} + 2^{2(S-2)} + \cdots + 1] \\
&= \mathbb{E}[\frac{2^{2S} - 1}{2^2 - 1}] \\
&= \frac{1}{2^2 - 1}\left(\left(\sum_{i=0}^{\infty}(1 - p_e)(p_e 2^2)^i\right) - 1\right).
\end{aligned}$$

Therefore, the estimation error is uniformly bounded if $p_e < \frac{1}{2^2}$.

(2) Necessity: For necessity, we use the fact that the disturbance $w[n - S]$ is independent of the non-erased observations present up to the time n. Therefore, the estimation error is lower bounded by

$$\begin{aligned}
\mathbb{E}[(x[n] - \mathbb{E}[x[n]|y^n])^2] &\geq \mathbb{E}[\mathbb{E}[(2^{S-1} w[n - S])^2|S]] \\
&= \mathbb{E}[2^{2(S-1)} \cdot \mathbf{1}(n - S \geq 0)] \\
&= \frac{1}{2^2}\left(\sum_{i=0}^{n}(1 - p_e)(p_e 2^2)^i\right)
\end{aligned}$$

Therefore, if $p_e \geq \frac{1}{2^2}$ the estimation error must diverge to ∞.

(3) Remarks: From the above proof, we can notice that the intermittent observability is decided by whether $p_e 2^2$ is less than 1. Here, 2 is the largest eigenvalue of the system, and p_e is the probability mass function (p.m.f.) tail of S which can be defined as $\exp \limsup_{s\to\infty} \frac{1}{s} \ln \mathbb{P}\{S = s\}$. Thus, we can think of two potential differences between scalar and vector systems: (i) The maximum eigenvalue (ii) The p.m.f. tail.

It turns out the latter is true, and the p.m.f. tail is the difference between scalar and vector systems. The following example shows why and how the p.m.f tail changes in vector systems.

2.5.1 Power Property

The power property answers question (b) of the previous section, "How can we solve each sub-problem?". Consider the example of [113].

$$\begin{cases}
\mathbf{x}[n+1] = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} \mathbf{x}[n] + \mathbf{w}[n] \\
y[n] = \beta[n] \begin{bmatrix} 1 & 1 \end{bmatrix} \mathbf{x}[n]
\end{cases}$$

Like above, we put $\mathbf{x}[0] = \mathbf{0}$, $\mathbf{w}[n]$ is 2-dimensional i.i.d. Gaussian vector with mean $\mathbf{0}$ and variance $\mathbf{I}$, and $\beta[n]$ is an independent Bernoulli process with probability $1-p_e$. We also extend the one-sided process to a two-sided process in the same way.

We can see the states are 2-dimensional, while the observations are 1-dimensional. Therefore, unlike scalar systems at least two observations are required to estimate the states. Moreover, if we write the observability Gramian matrix, we immediately notice cyclic behavior:

$$\begin{aligned}
&\mathbf{C} = \begin{bmatrix} 1 & 1 \end{bmatrix} \\
&\mathbf{CA^{-1}} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \\
&\mathbf{CA^{-2}} = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} \end{bmatrix} \\
&\mathbf{CA^{-3}} = \begin{bmatrix} \frac{1}{8} & -\frac{1}{8} \end{bmatrix} \\
&\vdots
\end{aligned}$$

Notice that $\mathbf{C}, \mathbf{CA^{-2}}, \mathbf{CA^{-4}}, \cdots$ are linearly dependent and $\mathbf{CA^{-1}}, \mathbf{CA^{-3}}, \mathbf{CA^{-5}}, \cdots$ are linearly dependent. Therefore, as observed in [113], we need both even and odd time observations to estimate the states. In this example, we will show that $p_e^{\star} = \frac{1}{2^4}$.

(1) Sufficiency: Let $p_e < \frac{1}{2^4}$. From (2.1) and (2.2), we can see that when $\beta[n-k] = 1$ the following equations hold:

$$\mathbf{x}[n] = \mathbf{A}^k \mathbf{x}[n-k] + \mathbf{A}^{k-1}\mathbf{w}[n-k] + \cdots + \mathbf{w}[n-1] \tag{2.14}$$

$$\begin{aligned}
\mathbf{y}[n-k] &= \mathbf{C}\mathbf{x}[n-k] + \mathbf{v}[n-k] \\
&= \mathbf{CA}^{-k}\mathbf{x}[n] - \underbrace{(\mathbf{CA}^{-1}\mathbf{w}[n-k] + \cdots + \mathbf{CA}^{-k}\mathbf{w}[n-1] - \mathbf{v}[n-k])}_{:=\mathbf{v}'[n-k]}
\end{aligned} \tag{2.15}$$

Here, we can see the variance of $\mathbf{v}'[n-k]$ is bounded as $\mathbb{E}[|\mathbf{v}'[n-k]|^2] = \mathbb{E}[(\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \mathbf{w}[n-1] + \cdots + \begin{bmatrix} \frac{1}{2^k} & \frac{1}{(-2)^k} \end{bmatrix} \mathbf{w}[n-k])^2] \leq 2\frac{\frac{1}{4}}{1-\frac{1}{4}} = \frac{2}{3}$.

Now, the stopping time S until we have enough observations to estimate the states becomes the first time until we get both even and odd time observations, i.e. $S := \inf\{k : 0 \leq k_1 < k_2 \leq k, \beta[n-k_1] = 1, \beta[n-k_2] = 1, k_1 \neq k_2 (\text{mod} 2)\}$. Here, the p.m.f. of S gets thicker than that of scalar cases. We can actually prove that the p.m.f. tail of S is $\exp \limsup_{s \to \infty} \frac{1}{s} \ln \mathbb{P}\{S = s\} = p_e^{\frac{1}{2}}$, which we will rigorously justify in Lemma 7.2. Thus, we can find $\delta, c > 0$ such that $p_e < \frac{1}{2^4} - \delta$ and $\mathbb{P}\{S = s\} \leq c\left(\frac{1}{2^4} - \delta\right)^{\frac{s}{2}}$ for all $s \in \mathbb{Z}^+$.

Now, we will analyze the performance of a suboptimal estimator which only uses two observations. Let $\widehat{\mathbf{x}}[n] := \begin{bmatrix} \mathbf{CA}^{-k_1} \\ \mathbf{CA}^{-k_2} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{y}[n-k_1] \\ \mathbf{y}[n-k_2] \end{bmatrix}$. Here, we can see the matrix inverse exists since k_1 and k_2 are even and odd time observations. Let $\mathcal{F}_\beta$ be the σ-field generated by $\beta[n]$. Then,

k_1, k_2, S are deterministic variables conditioned on $\mathcal{F}_\beta$. The estimation error is upper bounded by

$$\begin{aligned}
\mathbb{E}[|\mathbf{x}[n] - \widehat{\mathbf{x}}[n]|_2^2] &= \mathbb{E}[\mathbb{E}[|\mathbf{x}[n] - \widehat{\mathbf{x}}[n]|_2^2|\mathcal{F}_\beta]] = \mathbb{E}[\mathbb{E}[\left| \begin{bmatrix} \mathbf{CA}^{-k_1} \\ \mathbf{CA}^{-k_2} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{v}'[n-k_1] \\ \mathbf{v}'[n-k_2] \end{bmatrix} \right|_2^2 |\mathcal{F}_\beta]] \\
&\leq \mathbb{E}[\mathbb{E}[8 \cdot \left| \begin{bmatrix} \mathbf{CA}^{-k_1} \\ \mathbf{CA}^{-k_2} \end{bmatrix}^{-1} \right|_{max}^2 \cdot \left| \begin{bmatrix} \mathbf{v}'[n-k_1] \\ \mathbf{v}'[n-k_2] \end{bmatrix} \right|_{max}^2 |\mathcal{F}_\beta]] \\
&= 8 \cdot \mathbb{E}[\left| \begin{bmatrix} 2^{-k_1} & (-2)^{-k_1} \\ 2^{-k_2} & (-2)^{-k_2} \end{bmatrix}^{-1} \right|_{max}^2 \cdot \mathbb{E}[\left| \begin{bmatrix} \mathbf{v}'[n-k_1] \\ \mathbf{v}'[n-k_2] \end{bmatrix} \right|_{max}^2 |\mathcal{F}_\beta]] \\
&= 8 \cdot \mathbb{E}[\left| \frac{1}{2 \cdot 2^{-k_1} \cdot (-2)^{-k_2}} \begin{bmatrix} (-2)^{-k_2} & -(-2)^{-k_1} \\ -2^{-k_2} & 2^{-k_1} \end{bmatrix} \right|_{max}^2 \cdot \mathbb{E}[\left| \begin{bmatrix} \mathbf{v}'[n-k_1] \\ \mathbf{v}'[n-k_2] \end{bmatrix} \right|_{max}^2 |\mathcal{F}_\beta]] \\
&= 8 \cdot \mathbb{E}[\frac{1}{2^2} \left(\frac{2^{-k_1}}{2^{-k_1} \cdot 2^{-k_2}} \right)^2 \cdot \mathbb{E}[\left| \begin{bmatrix} \mathbf{v}'[n-k_1] \\ \mathbf{v}'[n-k_2] \end{bmatrix} \right|_{max}^2 |\mathcal{F}_\beta]] \\
&\leq 2 \cdot \mathbb{E}[2^{2k_2} \cdot \mathbb{E}[|\mathbf{v}'[n-k_1]|^2 + |\mathbf{v}'[n-k_2]|^2 |\mathcal{F}_\beta]] \\
&\leq \frac{8}{3}\mathbb{E}[2^{2S}] \leq \frac{8}{3} \sum_{s=0}^{\infty} 2^{2s} c \left(\frac{1}{2^4} - \delta \right)^{\frac{s}{2}} = \frac{8}{3} \sum_{s=0}^{\infty} c(1 - 2^4\delta)^{\frac{s}{2}} < \infty
\end{aligned}$$

Therefore, the estimation error is uniformly bounded for $p_e < \frac{1}{2^4}$.

(2) Necessity: We will show that the system is not intermittent observable when $p_e \geq \frac{1}{2^4}$. Denote the stopping time S' to be $\inf\{k \geq 0 : \beta[n-k] = 1, k \text{ is even}\}$. Then, $\mathbb{P}\{S' = 0\} = 1 - p_e$, $\mathbb{P}\{S' = 1\} = 0$, $\mathbb{P}\{S' = 2\} = (1 - p_e)p_e$, $\cdots$. Thus, the p.m.f. tail of S', $\exp \limsup_{s \to \infty} \frac{1}{s} \ln \mathbb{P}\{S' = s\}$, is $p_e^{\frac{1}{2}}$.

The state disturbance $\mathbf{w}[n - S']$ can be decomposed into two orthogonal components, $\mathbf{w}[n - S'] = \begin{bmatrix} 1 \\ 1 \end{bmatrix} w_1[n - S'] + \begin{bmatrix} 1 \\ -1 \end{bmatrix} w_2[n - S']$ where $w_1[n - S']$ and $w_2[n - S']$ are independent Gaussian random variables with zero mean and variance $\frac{1}{2}$. From the system equations (2.14), (2.15) and the definition of S', we can see that all the observations between time $n - S'$ and n are orthogonal to $w_2[n - S']$. Thus, the estimator does not know anything about $w_2[n - S']$ at time n, and thus we can lower bound the estimation error as follows.

$$\begin{aligned}
\mathbb{E}[(\mathbf{x}[n] - \mathbb{E}[\mathbf{x}[n]|\mathbf{y}^n])^2] &\geq \mathbb{E}[\mathbb{E}[|\mathbf{A}^{S'-1} \begin{bmatrix} 1 \\ -1 \end{bmatrix} w_2[n - S']|_2^2]|S'] \\
&\geq \mathbb{E}[2^{2(S'-1)}\mathbb{E}[(w_2[n - S'])^2|S']] = \frac{1}{2^3}\mathbb{E}[2^{2S'} \cdot \mathbf{1}(S' \geq n)] \\
&= \frac{1}{2^3} \sum_{i=0}^{\lfloor \frac{n}{2} \rfloor} (1 - p_e)(\sqrt{p_e} 2^2)^{2i}
\end{aligned}$$

Thus, if $p_e \geq \frac{1}{2^4}$ the estimation error diverges to ∞.

(3) Remarks: Compared to the scalar case, the p.m.f. tails of both S and S' in this vector system thicken to $\sqrt{p_e}$. This results in decreasing the critical erasure to $\frac{1}{2^4}$. The cyclic behavior of

the observability Gramian matrix, $\mathbf{C}$, $\mathbf{CA}^{-1}$, $\cdots$, causes the thickening of the p.m.f. tails. Thus, to capture this cyclic behavior of the observability Gramian matrix, we tentatively define an eigenvalue cycle as follows[9]: We say that the eigenvalues of $\mathbf{A}$, λ_1 and λ_2 belong to the same **eigenvalue cycle** if $\frac{\lambda_1}{\lambda_2}$ is a root of unity, i.e. $\left(\frac{\lambda_1}{\lambda_2}\right)^n = 1$ for some $n \in \mathbb{Z}$. Moreover, we say that $\mathbf{A}$ has **no eigenvalue cycles** if all the ratios between the eigenvalues of $\mathbf{A}$ are 1 or not roots of unity, which implies $\mathbf{A}$ has no nontrivial eigenvalue cycles.

To generalize this example and find the p.m.f. tail for arbitrary eigenvalue cycles, we use the idea of large deviations [24] which is equivalent to a union bound for simple cases. The idea goes as follows.

First, consider test channels that are erasure-type channels which would make the observability gramian rank-deficient. For this example, these would be the channel that erases every odd-time observations, the channel that erases every even-time observations and the channel that erases all observations.[10]

Next, measure the distance from the true channel to the test channels. In our case, the true channel is the channel without any restriction and the distance measure between the true and test channel is the hamming distance. For the test channels considered above, the distance to the odd-time erasure channel is 1 since we are restricting every one out of two indexes to be erasure. Likewise, the distance to the even-time erasure channel is 1 and the distance to the all erasure channel is 2.

Then, the large deviation principle intuitively says that the performance is decided by the minimum-distance test channel. For the example, the odd-time or even-time erasure channel whose distances are 1 will govern the performance.

So the effect of the eigenvalue cycle is to thicken the tail of the stopping time until you get enough observations to estimate the states. Analytically, the effect is equivalent to taking a proper power to the p_e and hence the name "power property".

2.5.2 Max Combining

This property answers the question (c) i.e. how we go from a single eigenvalue cycle to multiple eigenvalue cycles. Consider the following example with two eigenvalue cycles:

$$\begin{cases} \begin{bmatrix} x_1[n+1] \\ x_2[n+1] \\ x_3[n+1] \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} x_1[n] \\ x_2[n] \\ x_3[n] \end{bmatrix} + \mathbf{w}[n] \\ y[n] = \beta[n] \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \mathbf{x}[n] \end{cases}$$

As before, we let $\mathbf{x}[0] = \mathbf{0}$, $\mathbf{w}[n]$ be i.i.d. Gaussian with mean $\mathbf{0}$ and variance $\mathbf{I}$, and $\beta[n]$ be an independent Bernoulli process with probability $1 - p_e$. We also extend the one-sided process to a

[9] We will formally define eigenvalue cycles later in Section 2.6.

[10] In the actual characterization shown in Section 2.6, we will see that the set S' in (2.18) is a proxy for these test channels. This minimum distance to the test channels will be denoted as l_i in (2.18).

two-sided process. Here, we can see there are two eigenvalue cycles. One eigenvalue cycle is $\{2, -2\}$ and the other one is $\{3\}$, and these can be considered as two subsystems of the original system.

Then, from the previous arguments, we can see that the p.m.f. tails for these two systems are different. The p.m.f. tail for the eigenvalue cycle $\{3\}$ is p_e, while the p.m.f. tail for the eigenvalue cycle $\{2, -2\}$ is thickened to $p_e^{\frac{1}{2}}$. Therefore, the question is whether the thickened tail in the eigenvalue cycle $\{2, -2\}$ affects $\{3\}$. The answer turns out to be "No", and we can consider the two subsystems separately. Thus, in this example, the system is intermittent observable if and only if both subsystems are intermittent observable, i.e. $p_e^{\star} = \frac{1}{\max\{3^2, 2^{2\cdot 2}\}}$. The main idea to justify this is so-called *successive decoding* developed in information theory [21].

(1) Sufficiency: We will prove that $p_e < \frac{1}{\max\{3^2, 2^{2\cdot 2}\}}$ is sufficient for the intermittent observability using a successive decoding idea. The idea is simple. We first estimate the state $x_1[n]$. Then, since we have an estimate for $x_1[n]$, we can subtract the estimate from the system and reduce the dimension of the system. The remaining estimation error is considered as noise.

Let S be the stopping time until we receive three observations in the reverse process, i.e. $S := \inf\{k : 0 \leq k_1 < k_2 < k_3 \leq k, \beta[n-k_1] = 1, \beta[n-k_2] = 1, \beta[n-k_3] = 1\}$. Here, we can prove that the p.m.f. tail of S is the same as the scalar case. Therefore, $\exp \limsup_{s\to\infty} \ln \mathbb{P}\{S = s\} = p_e$, which we will justify in Lemma 7.2. Since we have the three observations at time $n - k_1$, $n - k_2$ and $n - k_3$, by the pigeon-hole principle at least two among them have to be congruent mod 2. Assume that k_1 and k_2 are both even. Then, by (2.15) we have

$$\begin{aligned}
y[n-k_1] &= \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{bmatrix}^{-k_1} \begin{bmatrix} x_1[n] \\ x_2[n] \\ x_3[n] \end{bmatrix} + v'[n-k_1] \\
&= \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 3^{-k_1} & 0 & 0 \\ 0 & 2^{-k_1} & 0 \\ 0 & 0 & 2^{-k_1} \end{bmatrix} \begin{bmatrix} x_1[n] \\ x_2[n] \\ x_3[n] \end{bmatrix} + v'[n-k_1] \\
&= \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}^{-k_1} \begin{bmatrix} x_1[n] \\ x_2[n] + x_3[n] \end{bmatrix} + v'[n-k_1]
\end{aligned}$$

Like in the above section, we can also prove that $\mathbb{E}[|v'[n-k]|^2] \leq 2\frac{\frac{1}{4}}{1-\frac{1}{4}} + \frac{\frac{1}{9}}{1-\frac{1}{9}} = \frac{19}{24}$. Here, we can notice that instantaneously at time $n - k_1$ and $n - k_2$ the system equation behaves like the following system with no eigenvalue cycles:

$$\begin{cases}
\begin{bmatrix} x_1[n+1] \\ x_2[n+1] + x_3[n+1] \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1[n] \\ x_2[n] + x_3[n] \end{bmatrix} + \begin{bmatrix} w_1[n] \\ w_2[n] + w_3[n] \end{bmatrix} \\
y[n] = \beta[n] \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x_1[n] \\ x_2[n] + x_3[n] \end{bmatrix}
\end{cases}$$

Consider the suboptimal estimator $\widehat{\mathbf{x}}[n] = \begin{bmatrix} \widehat{x}_1[n] \\ \widehat{x}_2[n] + \widehat{x}_3[n] \end{bmatrix} = \begin{bmatrix} 3^{-k_1} & 2^{-k_1} \\ 3^{-k_2} & 2^{-k_2} \end{bmatrix}^{-1} \begin{bmatrix} y[n-k_1] \\ y[n-k_2] \end{bmatrix}$. Let $\mathcal{F}_\beta$ be the σ-field generated by $\beta[n]$, and F be the event that k_1 and k_2 are even. The estimation error is upper bounded by

$$\mathbb{E}[\left| \begin{bmatrix} x_1[n] \\ x_2[n] + x_3[n] \end{bmatrix} - \widehat{\mathbf{x}}[n] \right|_2^2 |\mathcal{F}_\beta \cap F] = \mathbb{E}[\left| \begin{bmatrix} 3^{-k_1} & 2^{-k_1} \\ 3^{-k_2} & 2^{-k_2} \end{bmatrix}^{-1} \begin{bmatrix} v'[n-k_1] \\ v'[n-k_2] \end{bmatrix} \right|_2^2 |\mathcal{F}_\beta \cap F]$$

$$\leq 8 \cdot \left| \begin{bmatrix} 3^{-k_1} & 2^{-k_1} \\ 3^{-k_2} & 2^{-k_2} \end{bmatrix}^{-1} \right|_{max}^2 \cdot \mathbb{E}[\left| \begin{bmatrix} v'[n-k_1] \\ v'[n-k_2] \end{bmatrix} \right|_{max}^2 |\mathcal{F}_\beta \cap F]$$

$$= 8 \cdot \frac{19}{12} \cdot \left| \frac{1}{3^{-k_1}2^{-k_2} - 2^{-k_1}3^{-k_2}} \begin{bmatrix} 2^{-k_2} & -2^{-k_1} \\ -3^{-k_2} & 3^{-k_1} \end{bmatrix} \right|_{max}^2$$

$$= 8 \cdot \frac{19}{12} \cdot \left(\frac{2^{-k_1}}{3^{-k_1}2^{-k_2}\left(1 - \left(\frac{2}{3}\right)^{k_2-k_1}\right)} \right)^2$$

$$\leq 8 \cdot \frac{19}{12} \cdot 3^2 \cdot (3^{k_1} \cdot 2^{k_2-k_1})^2 \leq 57 \cdot 3^{2k_2} \leq 57 \cdot 3^{2S}$$

Likewise, we can prove the same bound holds even if k_1 and k_2 are not even. Therefore, the estimation error is bounded by $57 \cdot 3^{2S}$. Like the previous section, we can prove that if $p_e < \frac{1}{3^2}$ then $\mathbb{E}[3^{2S}] < \infty$. Thus, the expectation of the estimation error for $x_1[n]$ is uniformly bounded over time.

Once we estimate $x_3[n]$, we can subtract the estimation $\widehat{x}_3[n]$ from the observation, i.e. $y'[n] := y[n] - \beta[n]\widehat{x}_1[n]$. Then, the new system with the observation $y'[n]$ behaves like the following system:

$$\begin{cases} \begin{bmatrix} x_2[n+1] \\ x_3[n+1] \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} x_2[n] \\ x_3[n] \end{bmatrix} + \mathbf{w}[n] \\ y'[n] = \beta[n] \left(\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x_2[n] \\ x_3[n] \end{bmatrix} + (x_1[n] - \widehat{x}_1[n]) \right) \end{cases}$$

Since the expectation of the estimation error for $x_1[n]$ is uniformly bounded, it can be considered as a part of the observation noise.[11] In the same way as the previous section, we can prove that the estimation error for $x_2[n], x_3[n]$ is uniformly bounded if $p_e < \frac{1}{2^{2\cdot 2}}$. Notice that the minimum number of observations required to estimate the state by observability gramian matrix inversion is 3, the number of states. However, here we used more observations to apply successive decoding idea.

(2) Necessity: To prove that the example is not intermittent observable if $p_e \geq \frac{1}{\max\{3^2, 2^{2\cdot 2}\}}$, we will use a genie argument. If the states $x_2[n], x_3[n]$ are given to the estimator as side-information, the remaining system with $x_1[n]$ is a scalar system with the eigenvalue 3. We know that if $p_e \geq \frac{1}{3^2}$,

[11] Precisely speaking, the estimation error for $x_1[n]$ is a random variable which depends on the channel erasure process. Therefore, the rigorous proof of Section 2.8.3 has more steps to justify this argument.

$x_1[n]$ is not intermittent observable. We can also give $x_1[n]$ as side-information to conclude that $p_e \geq \frac{1}{2^{2\cdot 2}}$ is a necessary condition for intermittent observability.

(3) Remarks: In summary, we can solve problems with multiple eigenvalue cycles one by one without worrying about the existence of the other eigenvalue cycles. In other words, at each step we estimate the eigenvalue cycle associated with the largest eigenvalue. After the estimation, the eigenvalue cycle can be subtracted from the system except uniformly bounded estimation error. Then, we can simply repeat the steps for the remaining system. This procedure of successively solving and subtracting the unknowns is called successive decoding in information theory, and used as a decoding procedure for the multiple-access channel [21].

Therefore, we can conclude that the intermittent observability for a multiple eigenvalue-cycle system is bottlenecked by the hardest-to-estimate eigenvalue cycle, which manifests as the max operation in the critical erasure probability calculation.

2.5.3 Separability of Eigenvalue Cycles

The remaining question is what are the minimal irreducible sub-problems, whose answer can be expected to be eigenvalue cycles from the discussion up to now. In other words, we will understand general systems with multiple eigenvalue cycles by dividing into sub-systems with a single eigenvalue cycle. In the max-combining property, we already saw an example with multiple eigenvalue cycles. In the example, we first reduce the problem with multiple eigenvalue cycles to the problem with no eigenvalue cycles by sub-sampling plants. For example, in Section 2.5.2 we already saw that by sub-sampling (by 2), the system with an eigenvalue cycle (period 2) becomes a system with no eigenvalue cycles.

Thus, the question reduces to the fact that for systems with no eigenvalue cycles the critical erasure probability is $\frac{1}{|\lambda_{max}|^2}$, which will be shown in Corollary 2.1. To intuitively understand why this is true, we will consider three cases depending on the structure of $\mathbf{A}$.

The first case is when $\mathbf{A}$ is a diagonal matrix, and the magnitudes of its eigenvalues are distinct. In fact, this case is already proved in [113]. Let's consider a descriptive example when $\mathbf{A} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$, $\mathbf{C} = \begin{bmatrix} 1 & 1 \end{bmatrix}$. Then, the observability gramian of the system becomes $\begin{bmatrix} \mathbf{CA}^{n_1} \\ \mathbf{CA}^{n_2} \end{bmatrix} = \begin{bmatrix} 3^{n_1} & 2^{n_1} \\ 3^{n_2} & 2^{n_2} \end{bmatrix}$. To prove that the critical erasure probability is given as $\frac{1}{|\lambda_{max}|^2} = \frac{1}{3^2}$, it is enough to prove that the determinant of the observability gramian is large enough for almost all distinct n_1 and n_2. To justify this, we can use the fact that the ratio of the elements, $(\frac{3}{2})^n$, is an exponentially increasing function.

The second case is when $\mathbf{A}$ is a diagonal matrix, and the eigenvalues are distinct but have the same magnitude. Let's consider the system with $\mathbf{A} = \begin{bmatrix} e^{j} & 0 \\ 0 & e^{j\sqrt{2}} \end{bmatrix}$ and $\mathbf{C} = \begin{bmatrix} 1 & 1 \end{bmatrix}$. The observability gramian is given as $\begin{bmatrix}\mathbf{CA}^{n_1}\\ \mathbf{CA}^{n_2}\end{bmatrix} = \begin{bmatrix}e^{jn_1} & e^{j\sqrt{2}n_1}\\ e^{jn_2} & e^{j\sqrt{2}n_2}\end{bmatrix}$, and like above it is enough to show that the determinant of this observability gramian is large enough for almost all distinct n_1, n_2. Here, the arguments from [113] cannot work. For this, we instead used Weyl's criterion [54] which tells us that each element $(e^{jn}, e^{j\sqrt{2}n})$ behaves like a random variable $(e^{j\theta_1}, e^{j\theta_2})$ where θ_1 and θ_2 are independent random variables uniformly distributed on $[0, 2\pi]$. In fact, the effect of the hypothetical random variables $(e^{j\theta_1}, e^{j\theta_2})$ is quite similar to the actually randomly-dithered nonuniform sampling discussed in Section 2.7.

The last case is when $\mathbf{A}$ is a Jordan block matrix. Let's consider the system with $\mathbf{A} = \begin{bmatrix}2 & 1\\ 0 & 2\end{bmatrix}$ and $\mathbf{C} = \begin{bmatrix}1 & 0\end{bmatrix}$. The observability gramian is given as $\begin{bmatrix}\mathbf{CA}^{n_1}\\ \mathbf{CA}^{n_2}\end{bmatrix} = \begin{bmatrix}2^{n_1} & n_1 2^{n_1}\\ 2^{n_2} & n_2 2^{n_2}\end{bmatrix}$, and we have to show that the determinant of this observability gramian is large enough for almost all distinct n_1, n_2. Unlike the above cases, this example has polynomial terms in n_1, n_2. Exploiting this fact, we can reduce the problem to the fact that a polynomial function on n becomes zero only on a measure zero set.

By combining the insights from these three examples, we can prove that for a general matrix A with no eigenvalue cycles, the critical erasure probability is given as $\frac{1}{|\lambda_{max}|^2}$.

2.6 Intermittent Observability Characterization

Based on the intuition of the previous section, the intermittent observability condition can be characterized. We begin with the formal definition of a cycle.

Definition 2.14. *A multiset (a set that allows repetitions of its elements) $\{a_1, a_2, \cdots, a_l\}$ is called a cycle with length l and period p if $\left(\frac{a_i}{a_j}\right)^p = 1$ for all $i, j \in \{1, 2, \cdots, l\}$ and some $p \in \mathbb{N}$. Following convention, p is denoted*[12] *as*

$$p := \min\left\{n \in \mathbb{N} : \left(\frac{a_i}{a_j}\right)^n = 1, \forall i, j \in \{1, 2, \cdots, l\}\right\}.$$

For example, $\{a\}$ is a cycle with length 1 and period 1 by itself. $\{e^{j\omega}, e^{j(\omega + \frac{2\pi}{6})}\}$ is a cycle with length 2 and period 6. $\{e^{j}, e^{j\sqrt{2}}\}$ and $\{1, 2\}$ are not cycles. One trivially necessary condition for a_1, a_2 to belong to the same cycle is $|a_1| = |a_2|$. It can be also shown that cycles are closed under overlapping unions, meaning that if $\{a_1, a_2\}$ and $\{a_2, a_3\}$ are cycles, $\{a_1, a_2, a_3\}$ is also a cycle.

Now, we can define an eigenvalue cycle. It is well-known in linear system theory [17] that by properly changing coordinates, any linear system equations (2.1) can be written in an equivalent form with a Jordan matrix $\mathbf{A}$. Moreover, even though the MMSE value can be changed by a coordinate change, the condition for boundedness (stabilizability) remains the same. Rigorously, for any system matrix $\mathbf{A}$, there exists an invertible matrix $\mathbf{U}$ and an upper-triangular Jordan matrix

[12] We use $\frac{0}{0} = 1$, $\frac{1}{0} = \infty$, $1^{\infty} = \infty$ and $\frac{1}{\infty} = 0$.

$\mathbf{A}'$ such that $\mathbf{A} = \mathbf{U}\mathbf{A}'\mathbf{U}^{-1}$. We also define $\mathbf{B}' := \mathbf{U}\mathbf{B}$ and $\mathbf{C}' := \mathbf{C}\mathbf{U}$. Then, the matrix $\mathbf{A}'$ and $\mathbf{C}'$ can be written as the following form:

$$\begin{aligned}
&\mathbf{A}' = diag\{\mathbf{A_{1,1}}, \mathbf{A_{1,2}}, \cdots, \mathbf{A}_{\mu,\nu_\mu}\} \\
&\mathbf{C}' = \begin{bmatrix} \mathbf{C_{1,1}} & \mathbf{C_{1,2}} & \cdots & \mathbf{C}_{\mu,\nu_\mu} \end{bmatrix} \\
&\text{where} \\
&\quad \mathbf{A_{i,j}} \text{ is a Jordan block with an eigenvalue } \lambda_{i,j} \\
&\quad \{\lambda_{i,1}, \cdots, \lambda_{i,\nu_i}\} \text{ is a cycle with length } \nu_i \text{ and period } p_i \\
&\quad \text{For } i \neq i',\ \{\lambda_{i,j}, \lambda_{i',j'}\} \text{ is not a cycle} \\
&\quad \mathbf{C_{i,j}} \text{ is a } l \times \dim \mathbf{A_{i,j}} \text{ complex matrix.}
\end{aligned} \tag{2.16}$$

Since cycles are closed under overlapping unions, the eigenvalues of $\mathbf{A}$ can be uniquely partitioned into maximal cycles, $\{\lambda_{i,1}, \cdots, \lambda_{i,\nu_i}\}$. We call these cycles *eigenvalue cycles* and we say $\mathbf{A}$ has no eigenvalue cycle if all of its eigenvalue cycles are period 1.

Define

$$\begin{aligned}
&\mathbf{A_i} = diag\{\lambda_{i,1}, \cdots, \lambda_{i,\nu_i}\} \\
&\mathbf{C_i} = \begin{bmatrix} (\mathbf{C_{i,1}})_1 & \cdots & (\mathbf{C}_{\mathbf{i},\nu_\mathbf{i}})_1 \end{bmatrix} \\
&\text{where } (\mathbf{C_{i,j}})_1 \text{ is the first column of } \mathbf{C_{i,j}}.
\end{aligned} \tag{2.17}$$

In other words, we are dividing the original problem to sub-problems according to eigenvalue cycles.

Let l_i be the minimum cardinality among the sets $S' \subseteq \{0, 1, \cdots, p_i - 1\}$ whose resulting $S := \{0, 1, \cdots, p_i - 1\} \setminus S' = \{s_1, s_2, \cdots, s_{|S|}\}$ makes

$$\begin{bmatrix} \mathbf{C_i}\mathbf{A_i}^{s_1} \\ \mathbf{C_i}\mathbf{A_i}^{s_2} \\ \vdots \\ \mathbf{C_i}\mathbf{A_i}^{s_{|S|}} \end{bmatrix} \tag{2.18}$$

be rank deficient, i.e. the rank is strictly less than ν_i. Here, p_i and l_i will be used for the power property. l_i represents how many observations have to be erased out of p_i time steps to make the observability Gramian matrix rank deficient. This corresponds to the critical error event in large deviation theory.

Now, we can apply the max-combination property to characterize intermittent observability. Here is the main theorem of the chapter.

Theorem 2.7. *Given an intermittent system* $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \sigma, \sigma')$ *with probability of erasure* p_e*, let* $\sigma < \infty$*,* $\sigma' > 0$*, and* $(\mathbf{A}, \mathbf{B})$ *be controllable. Then, the intermittent system is intermittent observable* *if and only if*

$$p_e < \frac{1}{\max_{1 \le i \le \mu} |\lambda_{i,1}|^{2\frac{p_i}{l_i}}}.$$

or equivalently $\max_{1 \le i \le \mu} p_e^{\frac{l_i}{p_i}} |\lambda_{i,1}|^2 < 1$.

Proof. See Section 2.8.3 for sufficiency, and Section 2.8.4 for necessity. □

Here, we can notice that there is no assumption about stability or observability of the system. Let's first do a validity test of the theorem by trying stable modes and unobservable modes. If $|\lambda_{i,1}| < 1$, $\frac{1}{|\lambda_{i,1}|^{2\frac{p_i}{l_i}}} > 1$. Therefore, the stable modes do not contribute to the characterization of the critical erasure probability. If $(\mathbf{A_i}, \mathbf{C_i})$ are unobservable, $l_i = 0$. So, $\frac{1}{|\lambda_{i,1}|^{2\frac{p_i}{0}}} = 0$ if $|\lambda_{i,1}| \ge 1$ and $\frac{1}{|\lambda_{i,1}|^{2\frac{p_i}{0}}} = \infty$ if $|\lambda_{i,1}| < 1$. Therefore, if the unobservable modes are stable they do not affect the intermittent observability of the system and if they are not the system is not intermittent observable even if $p_e = 0$.

Even though in general l_i does not admit a closed form, it is computable for special cases.

Corollary 2.1. *Given an intermittent system* $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \sigma, \sigma')$ *with probability of erasure* p_e*, let* $\sigma < \infty$*,* $\sigma' > 0$*, and* $(\mathbf{A}, \mathbf{B})$ *be controllable. We further assume that* $(\mathbf{A}, \mathbf{C})$ *is observable and* $\mathbf{A}$ *has no eigenvalue cycles (i.e.* $\left(\frac{\lambda_i}{\lambda_j}\right)^n \ne 1$ *for all* $\lambda_i \ne \lambda_j$ *and* $n \in \mathbb{N}$*). Then, the intermittent system is intermittent observable if and only if* $p_e < \frac{1}{|\lambda_{max}|^2}$ *where* λ_{max} *is the largest magnitude eigenvalue of* $\mathbf{A}$*.*

Proof. Since $\mathbf{A}$ has no eigenvalue cycles, p_i equal to 1 for all i and $\mathbf{A_i}$ are scalars. Moreover, by the observability condition and Theorem 2.6, $\mathbf{C_i}$ is full-rank. Thus, $l_i = 1$ for all i and by Theorem 2.7 the critical erasure probability is $\frac{1}{\max_i |\lambda_{i,1}|^2} = \frac{1}{|\lambda_{max}|^2}$. □

For a more precise understanding of the critical erasure probability, we will focus on the case of a row vector $\mathbf{C}$ — i.e. single-output systems. Heuristically, a row vector $\mathbf{C}$ is the worst among $\mathbf{C}$ matrices since a vector observation is clearly better than a scalar observation.

Furthermore, we will also restrict the periods of the all eigenvalue cycles of $\mathbf{A}$ to be primes[13]. The technical reason for this restriction is that prime periods give us a useful invariance property of the sub-eigenvalue cycles. Let $\{\lambda_1, \lambda_2, \cdots, \lambda_l\}$ be an eigenvalue cycle with prime period p. Then, all subsets of $\{\lambda_1, \lambda_2, \cdots, \lambda_l\}$ with distinct elements are eigenvalue cycles with the same period p. This invariance property need not hold for eigenvalue cycles with composite periods as we will see by example later.

Corollary 2.2. *Given an intermittent system* $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \sigma, \sigma')$ *with probability of erasure* p_e*, let* $\sigma < \infty$*,* $\sigma' > 0$*, and* $(\mathbf{A}, \mathbf{B})$ *be controllable. We further assume that* $(\mathbf{A}, \mathbf{C})$ *is observable,* $\mathbf{C}$ *is a row*

[13]For convenience, we include 1 as a prime number here.

vector, and $\mathbf{A}$ *has only prime-period eigenvalue cycles of length* ν_i*. Then, the intermittent system is intermittent observable if and only if* $p_e < \frac{1}{\max_{1 \le i \le \mu} |\lambda_{i,1}|^{\frac{2p_i}{p_i - \nu_i + 1}}}$.

Proof. First, we introduce the following fact regarding Vandermonde matrix determinants [82]: Let p be a prime, $a_1, \cdots, a_n$ be pairwise incongruent in mod p and $b_1, \cdots, b_n$ be pairwise incongruent in mod p. Then,

$$\begin{bmatrix} e^{j2\pi\frac{a_1 b_1}{p}} & e^{j2\pi\frac{a_1 b_2}{p}} & \cdots & e^{j2\pi\frac{a_1 b_n}{p}} \\ e^{j2\pi\frac{a_2 b_1}{p}} & e^{j2\pi\frac{a_2 b_2}{p}} & \cdots & e^{j2\pi\frac{a_2 b_n}{p}} \\ \vdots & \vdots & \ddots & \vdots \\ e^{j2\pi\frac{a_n b_1}{p}} & e^{j2\pi\frac{a_n b_2}{p}} & \cdots & e^{j2\pi\frac{a_n b_n}{p}} \end{bmatrix}$$

is full rank. [82]

Furthermore, since $(\mathbf{A}, \mathbf{C})$ is observable and $\mathbf{C}$ is a row vector, by Theorem 2.6, $\lambda_{i,j}$ are distinct and $(\mathbf{C_{i,j}})_1$ are not zeros. Therefore, let $\lambda_{i,j} = |\lambda_i| e^{j2\pi\frac{q_{i,j}}{p_i}}$ where $q_{i,1}, \cdots, q_{i,\nu_i}$ are incongruent in mod p_i and p_i are primes.

Now, we will evaluate the critical erasure probability shown in Theorem 2.7. For this system, (2.18) can be written as

$$\begin{bmatrix} \mathbf{C_i}\mathbf{A_i}^{s_1} \\ \vdots \\ \mathbf{C_i}\mathbf{A_i}^{s_{|S|}} \end{bmatrix} = \begin{bmatrix} \lambda_{i,1}^{s_1} & \cdots & \lambda_{i,\nu_i}^{s_1} \\ \vdots & \ddots & \vdots \\ \lambda_{i,1}^{s_{|S|}} & \cdots & \lambda_{i,\nu_i}^{s_{|S|}} \end{bmatrix} \begin{bmatrix} (\mathbf{C_{i,1}})_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & (\mathbf{C}_{\mathbf{i},\nu_\mathbf{i}})_1 \end{bmatrix}$$

$$= \begin{bmatrix} |\lambda_i|^{s_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & |\lambda_i|^{s_{|S|}} \end{bmatrix} \begin{bmatrix} e^{j2\pi\frac{q_{i,1}}{p_i}s_1} & \cdots & e^{j2\pi\frac{q_{i,\nu_i}}{p_i}s_1} \\ \vdots & \ddots & \vdots \\ e^{j2\pi\frac{q_{i,1}}{p_i}s_{|S|}} & \cdots & e^{j2\pi\frac{q_{i,\nu_i}}{p_i}s_{|S|}} \end{bmatrix} \begin{bmatrix} (\mathbf{C_{i,1}})_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & (\mathbf{C}_{\mathbf{i},\nu_\mathbf{i}})_1 \end{bmatrix}$$

Since λ_i and $\mathbf{C_{i,j_1}}$ are non-zeros, the rank of $\begin{bmatrix} \mathbf{C_i}\mathbf{A_i}^{s_1} \\ \vdots \\ \mathbf{C_i}\mathbf{A_i}^{s_{|S|}} \end{bmatrix}$ is equal to the rank of

$$\begin{bmatrix} e^{j2\pi\frac{q_{i,1}}{p_i}s_1} & \cdots & e^{j2\pi\frac{q_{i,\nu_i}}{p_i}s_1} \\ \vdots & \ddots & \vdots \\ e^{j2\pi\frac{q_{i,1}}{p_i}s_{|S|}} & \cdots & e^{j2\pi\frac{q_{i,\nu_i}}{p_i}s_{|S|}} \end{bmatrix}.$$

Furthermore, since $q_{i,1}, \cdots, q_{i,\nu_1}$ are incongruent in mod p_i and $s_1, \cdots, s_{|S|}$ are also incongruent in mod p_i, by the property of the Vandermonde matrix discussed above, the rank of the observability gramian is greater or equal to ν_i if and only if $|S| \ge \nu_i$.

Therefore, l_i of (2.18) is $p_i - \nu_i + 1$, and the corollary follows from Theorem 2.7. □

One may wonder why we could not get a simple answer in Theorem 2.7 unlike Corollary 2.2. To understand this, consider two potential extensions of Corollary 2.2:

(1) Eigenvalue cycles with periods that are composite numbers:

Consider $\mathbf{A} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2e^{j\frac{2\pi}{16}} & 0 \\ 0 & 0 & 2e^{j\frac{2\pi}{16}9} \end{bmatrix}$ and $\mathbf{C} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$. The eigenvalue cycle has length 3 and period 16. If we naively apply the formula of Corollary 2.2 then we would get a critical value $\frac{1}{2^{2\cdot\frac{16}{16-3+1}}} = \frac{1}{2^{\frac{16}{7}}}$. However, if we consider the sub-eigenvalue cycle $\{2e^{j\frac{2\pi}{16}}, 2e^{j\frac{2\pi}{16}9}\}$, the length is 2 and the period is 2. The formula of Corollary 2.2 gives $\frac{1}{2^{2\cdot\frac{2}{2-2+1}}} = \frac{1}{2^4}$ as a critical value, which gives a tighter condition than the previous one. In fact, the latter value is the correct critical erasure probability. Because the period invariant property does not hold for a composite number cycle, the longest cycle does not necessarily give the right critical probability.

(2) A general matrix $\mathbf{C}$, multiple-output systems: If we have a vector observation, an eigenvalue cycle can be divided into smaller cycles. As an extreme case, when $\mathbf{C}$ is an identity matrix every eigenvalue cycle is divided into trivial cycles with length 1 and the critical erasure probability becomes $\frac{1}{|\lambda_{max}|^2}$ as observed in [95]. Consider now $\mathbf{A} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2e^{j\frac{2\pi}{5}} & 0 & 0 \\ 0 & 0 & 2e^{j\frac{2\pi}{5}2} & 0 \\ 0 & 0 & 0 & 2e^{j\frac{2\pi}{5}3} \end{bmatrix}$ and $\mathbf{C} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & \delta \end{bmatrix}$. The eigenvalue cycle $\{2, 2e^{j\frac{2\pi}{5}}, 2e^{j\frac{2\pi}{5}2}, 2e^{j\frac{2\pi}{5}3}\}$ of $\mathbf{A}$ has length 4 and period 5. However, if $\delta \neq 0$, by elementary row operations $\mathbf{C}$ can be converted to $\begin{bmatrix} 1 & 2 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$. Thus, the eigenvalue cycle is divided into two sub-cycles, $\{2, 2e^{\frac{2\pi}{5}}, 2e^{\frac{2\pi}{5}2}\}$ and $\{2e^{\frac{2\pi}{5}3}\}$. The longer cycle with length 3 would dominate and the critical erasure probability would be $\frac{1}{2^{2\cdot\frac{5}{5-3+1}}} = \frac{1}{2^{\frac{10}{3}}}$. Meanwhile, if $\delta = 0$, the second row of $\mathbf{C}$ would be ignorable. Thus, the eigenvalue cycle would not be divided and the critical erasure probability would be $\frac{1}{2^{2\cdot\frac{5}{5-4+1}}} = \frac{1}{2^{\frac{10}{2}}}$.

In this example, we can see that the critical erasure probability depends on whether δ is equal to 0 or not, which is related to the rank of $\mathbf{C}$. Thus, it is inevitable to have a rank condition of some sort in the characterization of the critical erasure probability.

2.6.1 Extension to Intermittent Kalman Filtering with Parallel Channels

The concept of eigenvalue cycles and the divide-and-conquer approach can be also applied to extensions and variations of the intermittent Kalman filtering.

Let's consider intermittent Kalman filtering with parallel erasure channels as introduced in [33].

$$\mathbf{x}[n+1] = \mathbf{A}\mathbf{x}[n] + \mathbf{B}\mathbf{w}[n]$$
$$\mathbf{y_1}[n] = \beta_1[n](\mathbf{C_1}\mathbf{x}[n] + \mathbf{v_1}[n])$$
$$\vdots$$
$$\mathbf{y_d}[n] = \beta_d[n](\mathbf{C_d}\mathbf{x}[n] + \mathbf{v_d}[n])$$

Here n is the non-negative integer-valued time index, and $\mathbf{x}[n] \in \mathbb{C}^m$, $\mathbf{w}[n] \in \mathbb{C}^g$, $\mathbf{y_i}[n] \in \mathbb{C}^{l_i}$, $\mathbf{v_i}[n] \in \mathbb{C}^{l_i}$, $\mathbf{A} \in \mathbb{C}^{m\times m}$, $\mathbf{B} \in \mathbb{C}^{m\times g}$, $\mathbf{C_i} \in \mathbb{C}^{l_i\times m}$. The underlying randomness comes from $\mathbf{x}[0]$, $\mathbf{w}[n]$, $\mathbf{v_i}[n]$ and $\beta_i[n]$. $\mathbf{x}[0]$, $\mathbf{w}[n]$ and $\mathbf{v_i}[n]$ are independent Gaussian vectors with zero mean, and there exist positive σ^2 and σ'^2 such that

$$\mathbb{E}[\mathbf{x}[0]\mathbf{x}[0]^\dagger] \preceq \sigma^2\mathbf{I}$$
$$\mathbb{E}[\mathbf{w}[n]\mathbf{w}[n]^\dagger] \preceq \sigma^2\mathbf{I}$$
$$\mathbb{E}[\mathbf{v_i}[n]\mathbf{v_i}[n]^\dagger] \preceq \sigma^2\mathbf{I}$$
$$\mathbb{E}[\mathbf{w}[n]\mathbf{w}[n]^\dagger] \succeq \sigma'^2\mathbf{I}$$
$$\mathbb{E}[\mathbf{v_i}[n]\mathbf{v_i}[n]^\dagger] \succeq \sigma'^2\mathbf{I}.$$

$\beta_i[n]$ are independent Bernoulli random processes with erasure probabilities $p_{e,i}$.

We call this system as an intermittent system $(\mathbf{A}, \mathbf{B}, \mathbf{C_i})$ with erasure probabilities $p_{e,i}$.

Since the observations go through independent parallel erasure channels, we can expect diversity gain [99], i.e. even though the observations from some channels are lost, we can still estimate the state based on other successfully transmitted observations. At the first glance, this extension may seem much harder than the original problem since we have to characterize the whole region $(p_{e,1}, \cdots, p_{e,d})$ rather than a single critical erasure value. However, a simple extension of Theorem 2.7 turns out to be enough to characterize this critical erasure probability region. As in Section 2.6, let $\mathbf{A} = \mathbf{U}\mathbf{A}'\mathbf{U}^{-1}$ where $\mathbf{U}$ is an invertible matrix and $\mathbf{A}'$ is an upper-triangular Jordan matrix. We also define $\mathbf{B}' := \mathbf{U}\mathbf{B}$ and $\mathbf{C_i'} := \mathbf{C_i}\mathbf{U}$.

Then, we can make the following generalized definitions of (2.16), (2.17), (2.18) for $\mathbf{A}'$ and $\mathbf{C_i'}$.

$$\mathbf{A}' = diag\{\mathbf{A_{1,1}}, \mathbf{A_{1,2}}, \cdots, \mathbf{A}_{\mu,\nu_\mu}\}$$
$$\mathbf{C_i'} = \begin{bmatrix} \mathbf{C_{1,1,i}} & \mathbf{C_{1,2,i}} & \cdots & \mathbf{C}_{\mu,\nu_\mu,\mathbf{i}} \end{bmatrix}$$

where

$\mathbf{A_{i,j}}$ is a Jordan block matrix with an eigenvalue $\lambda_{i,j}$

$\{\lambda_{i,1}, \cdots, \lambda_{i,\nu_i}\}$ is a cycle with length ν_i and period p_i

For $i \neq i'$, $\{\lambda_{i,j}, \lambda_{i',j'}\}$ is not a cycle

$\mathbf{C_{i,j,k}}$ is a $l_k \times \dim \mathbf{A_{i,j}}$ matrix.

Denote

$$\mathbf{A_i} = diag\{\lambda_{i,1}, \cdots, \lambda_{i,\nu_i}\}$$
$$\mathbf{C_{i,j}} = \left[(\mathbf{C_{i,1,j}})_1, \cdots, (\mathbf{C_{i,\nu_i,j}})_1\right]$$
where $(\mathbf{C_{i,j,k}})_1$ is the first column of $\mathbf{C_{i,j,k}}$.

Let $(l_{i,1}, l_{i,2}, \cdots, l_{i,d})$ be the cardinality vector of the sets $S'_1, S'_2, \cdots, S'_d$ such that $S_j := \{0, 1, \cdots, p_i - 1\} \setminus S'_j = \{s_{j,1}, s_{j,2}, \cdots, s_{j,|S_j|}\}$ and

$$\begin{bmatrix} \mathbf{C_{i,1}}\mathbf{A_i}^{s_{1,1}} \\ \vdots \\ \mathbf{C_{i,1}}\mathbf{A_i}^{s_{1,|S_1|}} \\ \mathbf{C_{i,2}}\mathbf{A_i}^{s_{2,1}} \\ \vdots \\ \mathbf{C_{i,d}}\mathbf{A_i}^{s_{d,|S_d|}} \end{bmatrix}$$

is rank deficient, i.e. has rank strictly less than ν_i. Denote L_i as a set of all such vectors.

Then, intermittent observability with parallel channels is characterized as follows.

Proposition 2.1. *Given an intermittent system $(\mathbf{A}, \mathbf{B}, \mathbf{C_i}, \sigma, \sigma')$ with probabilities of erasures $(p_{e,1}, \cdots, p_{e,d})$, let $\sigma < \infty$, $\sigma' > 0$, and $(\mathbf{A}, \mathbf{B})$ be controllable. Then, the intermittent system is intermittent observable if and only if*

$$\max_{1 \le i \le \mu} \max_{(l_{i,1}, l_{i,2}, \cdots, l_{i,d}) \in L_i} \left(\prod_{1 \le j \le d} p_{e,j}^{\frac{l_{i,j}}{p_i}} \right) |\lambda_{i,1}|^2 < 1.$$

We omit the proof of the proposition, since it is similar to that of Theorem 2.7.

Compared to Theorem 2.7, the max-combination and separability principle remain the same, but the test channels in the power property become more complicated. Here, $(S'_1, \cdots, S'_d)$ represents the test channels such that when they are erased, the observability Gramian becomes rank-deficient. $(l_{i,1}, \cdots, l_{i,d})$ represents the distance vector to these test channels.

2.7 Intermittent Kalman Filtering with Nonuniform Sampling

In the previous section, we proved that eigenvalue cycles are the only factor that prevents us from having the critical erasure probability be $\frac{1}{|\lambda_{max}|^2}$. Based on this understanding, we can look for a simple way to avoid this troublesome phenomenon. Here, we propose nonuniform sampling as a simple way of breaking the eigenvalue cycles and achieving the critical value $\frac{1}{|\lambda_{max}|^2}$.

As an intuitive example, consider $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Then, $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, $\mathbf{A}^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\mathbf{A}^3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, $\mathbf{A}^4 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\cdots$. What the eigenvalue cycle is capturing is that half of $\mathbf{A}, \mathbf{A}^2, \mathbf{A}^3, \cdots$ are identical. Therefore, the question is how we can make every matrix in $\mathbf{A}, \mathbf{A}^2, \mathbf{A}^3, \cdots$ distinct. To simplify the question, consider the sequence of $-1, 1, -1, 1, \cdots$ which corresponds to $(2,2)$ elements of $\mathbf{A}, \mathbf{A}^2, \mathbf{A}^3, \cdots$.

Rewrite this sequence $-1, 1, -1, 1, \cdots$ as $(e^{j\pi})^1, (e^{j\pi})^2, (e^{j\pi})^3, (e^{j\pi})^4, \cdots$ and introduce a jitter t_i to each sampling time. The resulting sequence becomes $(e^{j\pi})^{1+t_1}, (e^{j\pi})^{2+t_2}, (e^{j\pi})^{3+t_3}, (e^{j\pi})^{4+t_4}, \cdots$ and if t_is are uniformly distributed i.i.d. random variables on $[0, T]$ each element in the sequence is distinct almost surely as long as $T > 0$.

Operationally, this idea can be implemented as follows: at design-time, the sensor and the estimator agree on the nonuniform sampling pattern which is a realization of i.i.d. random variables whose distribution is uniform on $[0, T]$ ($T > 0$). Whenever the sensor samples the system, it jitters its sampling time according to this nonuniform pattern. Knowing the sampling time jitter, the sampled continuous-time system looks like a discrete *time-varying* system to the estimator. The joint Gaussianity between the observation and the state is preserved, and furthermore, Kalman filters are optimal even for time-varying systems! This intermittent Kalman filtering problem with nonuniform samples has the critical erasure probability $\frac{1}{|\lambda_{max}|^2}$ almost surely. Therefore, an eigenvalue cycle is breakable by nonuniform sampling.

One may be bothered by the probabilistic argument on the nonuniform sampling pattern. However, this probabilistic proof is an indirect argument for the existence of an appropriate deterministic nonuniform sampling pattern, which is similar to how the existence of capacity-achieving codes is proved in information theory [93].

To write the scheme formally, consider a continuous-time dynamic system:

$$d\mathbf{x_c}(t) = \mathbf{A_c}\mathbf{x_c}(t)dt + \mathbf{B_c}d\mathbf{W_c}(t) \tag{2.19}$$

$$\mathbf{y_c}(t) = \mathbf{C_c}\mathbf{x_c}(t) + \mathbf{D_c}\frac{d\mathbf{V_c}(t)}{dt}. \tag{2.20}$$

Here t is the non-negative real-valued time index. $\mathbf{W_c}(t)$ and $\mathbf{V_c}(t)$ are independent g and l-dimension standard Wiener processes respectively, i.e. for $a, b \geq 0$, $\mathbf{W_c}(a+b) - \mathbf{W_c}(b)$ is distributed as $\mathcal{N}(\mathbf{0}, a\mathbf{I})$ and $\mathbf{V_c}(a+b) - \mathbf{V_c}(b)$ is also distributed as $\mathcal{N}(0, a\mathbf{I})$. $\mathbf{A_c} \in \mathbb{C}^{m\times m}$, $\mathbf{B_c} \in \mathbb{C}^{m\times g}$, $\mathbf{C_c} \in \mathbb{C}^{l\times m}$, and $\mathbf{D_c} \in \mathbb{C}^{l\times l}$ where $\mathbf{D_c}$ is invertible. Thus, $\mathbf{x}[n] \in \mathbb{C}^m$ and $\mathbf{y}[n] \in \mathbb{C}^l$. For convenience, we assume $\mathbf{x}[0] = 0$ but the results of this chapter hold for any $\mathbf{x}[0]$ with finite variance. Throughout this chapter, we use the Ito's integral [32, p.80] for stochastic calculus.

The process of (2.19) is known as Ornstein-Uhlenbeck process [32, p.109] whose solution is

$\mathbf{x_c}(t) = e^{\mathbf{A_c}t}\mathbf{x_c}(0) + \int_0^t e^{\mathbf{A_c}(t-t')}\mathbf{B_c}d\mathbf{W_c}(t')$. Therefore, for $t_1 \leq t_2$ we have

$$\begin{aligned}
\mathbf{x_c}(t_2) &= e^{\mathbf{A_c}t_2}\mathbf{x_c}(0) + \int_0^{t_2} e^{\mathbf{A_c}(t_2-t')}\mathbf{B_c}d\mathbf{W_c}(t') \\
&= e^{\mathbf{A_c}(t_2-t_1)}\left(e^{\mathbf{A_c}t_1}\mathbf{x_c}(0) + \int_0^{t_2} e^{\mathbf{A_c}(t_1-t')}\mathbf{B_c}d\mathbf{W_c}(t')\right) \\
&= e^{\mathbf{A_c}(t_2-t_1)}\left(e^{\mathbf{A_c}t_1}\mathbf{x_c}(0) + \int_0^{t_1} e^{\mathbf{A_c}(t_1-t')}\mathbf{B_c}d\mathbf{W_c}(t') + \int_{t_1}^{t_2} e^{\mathbf{A_c}(t_1-t')}\mathbf{B_c}d\mathbf{W_c}(t')\right) \\
&= e^{\mathbf{A_c}(t_2-t_1)}\left(\mathbf{x_c}(t_1) + \int_{t_1}^{t_2} e^{\mathbf{A_c}(t_1-t')}\mathbf{B_c}d\mathbf{W_c}(t')\right)
\end{aligned} \tag{2.21}$$

which can be rewritten as

$$\mathbf{x_c}(t_1) = e^{\mathbf{A_c}(t_1-t_2)}\mathbf{x_c}(t_2) - \int_{t_1}^{t_2} e^{\mathbf{A_c}(t_1-t')}\mathbf{B_c}d\mathbf{W_c}(t'). \tag{2.22}$$

The point of doing this is to understand the values of the states during sampling intervals in terms of the states at the end of the interval.

Let's say we want to sample the system with a sampling interval I ($I > 0$). Conventional samplers uses integration filters to sample, i.e. in the uniform sampling case, the nth sample $\mathbf{y}[n]$ corresponds to the integration of $\mathbf{y_c}(t)$ for $(n-1)I \leq t < nI$:

$$\mathbf{y}[n] = \int_{(n-1)I}^{nI} \mathbf{y_c}(t)dt.$$

Nonuniform sampling can be thought of in two ways with respect to sampler's integration filters: (1) The starting times of the integrations are uniform but the sampling intervals are non-uniform. (2) The sampling intervals are uniform, but the starting times of the integrations are non-uniform. Since the analysis and performance is similar in both cases, we will focus on the latter case. To take the nth sample of the system, the non-uniform sampler takes the integration of $\mathbf{y_c}(t)$

for $(n-1)I - t_n \leq t < nI - t_n$:

$$\begin{aligned}
\mathbf{y_o}[n] &= \int_{(n-1)I-t_n}^{nI-t_n} \mathbf{y_c}(t)dt \\
&= \int_{(n-1)I-t_n}^{nI-t_n} \mathbf{C_c}\mathbf{x_c}(t)dt + \int_{(n-1)I-t_n}^{nI-t_n} \mathbf{D_c}d\mathbf{V_c}(t) \quad (2.23)\\
&= \int_{(n-1)I-t_n}^{nI-t_n} \mathbf{C_c}\left(e^{\mathbf{A_c}(t-(nI-t_n))}\mathbf{x_c}(nI-t_n) - \int_{t}^{nI-t_n} e^{\mathbf{A_c}(t-t')}\mathbf{B_c}d\mathbf{W_c}(t')\right)dt \\
&+ \int_{(n-1)I-t_n}^{nI-t_n} \mathbf{D_c}d\mathbf{V_c}(t) \quad (2.24)\\
&= \left(\int_{(n-1)I-t_n}^{nI-t_n} \mathbf{C_c}e^{\mathbf{A_c}(t-(nI-t_n))}dt\right)\mathbf{x_c}(nI-t_n) \\
&- \int_{(n-1)I-t_n}^{nI-t_n}\int_{t}^{nI-t_n} \mathbf{C_c}e^{\mathbf{A_c}(t-t')}\mathbf{B_c}d\mathbf{W_c}(t')dt + \int_{(n-1)I-t_n}^{nI-t_n} \mathbf{D_c}d\mathbf{V_c}(t) \\
&= \underbrace{\left(\int_0^I \mathbf{C_c}e^{\mathbf{A_c}(t-I)}dt\right)}_{:=\mathbf{C}}\mathbf{x_c}(nI-t_n) \\
&\underbrace{- \int_{(n-1)I-t_n}^{nI-t_n}\int_{t}^{nI-t_n} \mathbf{C_c}e^{\mathbf{A_c}(t-t')}\mathbf{B_c}d\mathbf{W_c}(t')dt + \int_{(n-1)I-t_n}^{nI-t_n} \mathbf{D_c}d\mathbf{V_c}(t)}_{:=\mathbf{v}[n]} \quad (2.25)
\end{aligned}$$

Here (2.23) follows from (2.20), and (2.24) follows from (2.22). Since $\mathbf{y_o}[n]$ is transmitted over the erasure channel, the intermittent system $(\mathbf{A_c}, \mathbf{B_c}, \mathbf{C})$ with nonuniform samples and erasure probability p_e has the following system equation:

$$d\mathbf{x_c}(t) = \mathbf{A_c}\mathbf{x_c}(t)dt + \mathbf{B_c}d\mathbf{W_c}(t) \quad (2.26)$$

$$\mathbf{y}[n] = \beta[n](\mathbf{C}\mathbf{x_c}(nI-t_n) + \mathbf{v}[n]) \quad (2.27)$$

where $\mathbf{y}[n] \in \mathbb{C}^l$ and $\beta[n]$ is an independent Bernoulli random process with erasure probability p_e. The variance of $\mathbf{v}[n]$ is uniformly bounded since the integration interval is bounded, but $\mathbf{v}[n]$ can be correlated since the integration intervals could overlap. Since $\mathbf{C}$ is a function of $\mathbf{C_c}$, the observability of $(\mathbf{A_c}, \mathbf{C_c})$ does not necessarily imply the observability of $(\mathbf{A_c}, \mathbf{C})$ while the observability of $(\mathbf{A_c}, \mathbf{C})$ always implies the observability of $(\mathbf{A_c}, \mathbf{C_c})$.

Figure 2.3 shows the system diagram for intermittent Kalman filtering with nonuniform sampling. The nonuniform sampler samples the plant according to the nonuniform sampling pattern t_n and generates observations $y_o[n]$. The observation is transmitted through the real erasure channel without any coding. Then, the estimator tries to estimate the state $x_c(t)$ based on its received signals y^n and the nonuniform sampling pattern t^n.

As before, the intermittent system $(\mathbf{A_c}, \mathbf{B_c}, \mathbf{C})$ with nonuniform samples is called intermittent observable if there exists a causal estimator $\widehat{\mathbf{x}}(t)$ of $\mathbf{x}(t)$ based on $\mathbf{y}[\lfloor \frac{t}{I} \rfloor], \cdots, \mathbf{y}[0]$ such

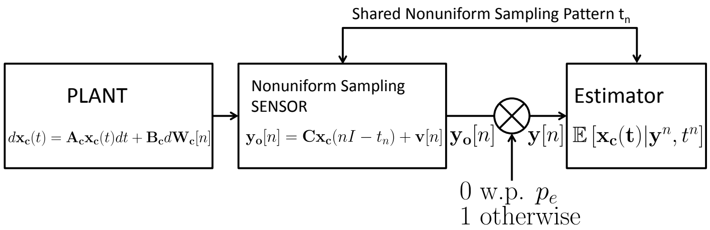


Figure 2.3: System diagram for 'intermittent Kalman filtering with nonuniform sampling'. The sensor samples the plant according to the nonuniform sampling pattern t_n, and sends the observation through the real erasure channel without any coding. The estimator tries to estimate the state based on its received signals and the nonuniform sampling pattern t_n.

that

$$\sup_{t \in \mathbb{R}^+} \mathbb{E}[(\mathbf{x}(t) - \widehat{\mathbf{x}}(t))^\dagger(\mathbf{x}(t) - \widehat{\mathbf{x}}(t))] < \infty.$$

Intermittent observability with nonuniform samples is characterized by the following theorem.

Theorem 2.8. *Let t_n be i.i.d. random variables uniformly distributed on $[0, T]$ ($T > 0$), and $(\mathbf{A_c}, \mathbf{B_c})$ be controllable. When $(\mathbf{A_c}, \mathbf{C})$ has unobservable and unstable eigenvalues — i.e. $\exists \lambda \in \mathbb{C}^+$ such that $\begin{bmatrix} \lambda\mathbf{I} - \mathbf{A_c} \\ \mathbf{C} \end{bmatrix}$ is rank deficient —, the intermittent system $(\mathbf{A_c}, \mathbf{B_c}, \mathbf{C})$ with nonuniform samples is not intermittent observable for all p_e. Otherwise, the intermittent system $(\mathbf{A_c}, \mathbf{B_c}, \mathbf{C})$ with nonuniform samples is intermittent observable if and only if $p_e < \frac{1}{|e^{2\lambda_{max} I}|}$. Here λ_{max} is the eigenvalue of $\mathbf{A_c}$ with the largest real part.*

Proof. See Section 2.8.1 for sufficiency, and Section 2.8.2 for necessity. □

Since $\exp((\text{eigenvalue of } \mathbf{A_c})I)$ corresponds to the eigenvalue of the sampled discrete time system, the critical value of Theorem 2.8 is equivalent to that of Corollary 2.1. The nonuniform sampling allows us to no longer care if eigenvalue cycles could exist for the original continuous-time system under uniform sampling.

Nonuniform sampling is the right way of breaking eigenvalue cycles from a practical point of view. The critical erasure probability of $\frac{1}{|\lambda_{max}|^2}$ can thus be achieved not only by using the computationally challenging estimation-before-packetization strategy of [87], but also by the simple memoryless approach of dithered sampling before packetization. And so, even if the sensors were themselves distributed, the critical erasure probability with nonuniform sampling is still *critical value optimal* in a sense that they can achieve the same critical erasure probability as sensors with causal or noncausal information about the erasure pattern and with unbounded complexity.

2.7.1 Extensions of Intermittent Kalman Filtering with Nonuniform Sampling

In this section, we discuss variations and extensions of intermittent Kalman filtering with nonuniform samples. Since the proofs of the results shown in this section are similar to that of Theorem 2.8, we only present the results without proofs.

General Distribution on t_n

First, we relax the condition on the distribution of t_n of Theorem 2.8. There, we assume that t_n are identically and uniformly distributed. However, they do not have to be identical or uniform.

Proposition 2.2. *Assume that $t_0, t_1, \cdots$ are independent and there exist $a, c > 0$ such that $\mathbb{P}\{|t_n| \geq a\} = 0$ and $\mathbb{P}\{t_n \in B\} \leq c|B|_{\mathcal{L}}$ for all $n \in \mathbb{Z}^+$ and $B \in \mathcal{B}$, where $\mathcal{B}$ is Borel σ-algebra and $|\cdot|_{\mathcal{L}}$ is Lebesgue measure. Then, Theorem 2.8 still holds, i.e. if $(\mathbf{A_c}, \mathbf{C})$ has no unobservable and unstable eigenvalues, the intermittent system with nonuniform samples is intermittent observable if and only if $p_e < \frac{1}{|e^{2\lambda_{max}I}|}$.*

For the proof of the proposition, we can repeat the proof steps of Theorem 2.8 using an improper distribution μ such that $\mu(A) = c|A \cap [-a, a]|_{\mathcal{L}}$.

Deterministic Sequences for t_n

The randomness assumption on t_n can be also removed. As we mentioned earlier, the probabilistic proof is an indirect proof for the existence of deterministic nonuniform sampling patterns. In fact, any nonuniform sequence satisfying Weyl's criteria —which gives the sufficient and necessary condition for a sequence equidistributed on the interval — can be used to break eigenvalue cycles.

Proposition 2.3. *Let a sequence $t_n \in [0, T]$ satisfy Weyl's criteria, i.e. for all $h \in \mathbb{Z} \setminus \{0\}$, $\lim_{N \to \infty} |\frac{1}{N} \sum_{1 \leq n \leq N} e^{j2\pi h \cdot \frac{t}{T}}| = 0$. Then, Theorem 2.8 still holds, i.e. if $(\mathbf{A_c}, \mathbf{C})$ has no unobservable and unstable eigenvalues, the intermittent system with nonuniform samples is intermittent observable if and only if $p_e < \frac{1}{|e^{2\lambda_{max}I}|}$.*

For example, a sequence like $t_n = \sqrt{2}n - \lfloor \sqrt{2}n \rfloor$ can be used to break eigenvalue cycles. The proof is by merging the proofs of Theorem 2.7 and Theorem 2.8.

Nonuniform-length integration interval

In Theorem 2.8, we introduce nonuniform sampling by changing the starting time of the length of the integration. Another way of introducing nonuniform sampling is changing the integration interval. To take the nth sample of the system, the sensor integrates $\mathbf{y_c}(t)$ from $(n-1)I - t_n$

to nI. Parallel to (2.25), we have the following equation.

$$\begin{aligned}
\mathbf{y_o}[n] &= \int_{(n-1)I-t_n}^{nI} \mathbf{y_c}(t)dt \\
&= \left(\int_{(n-1)I-t_n}^{nI} \mathbf{C_c} e^{\mathbf{A_c}(t-nI)}\right)\mathbf{x_c}(nI) \\
&- \int_{(n-1)I-t_n}^{nI}\int_{t}^{nI-t_n} \mathbf{C_c} e^{\mathbf{A_c}(t-t')}\mathbf{B_c} d\mathbf{W_c}(t')dt + \int_{(n-1)I-t_n}^{nI} \mathbf{D_c} d\mathbf{V_c}(t) \\
&= \underbrace{\left(\int_{0}^{n+t_n} \mathbf{C_c} e^{\mathbf{A_c}(t-nI-t_n)}\right)}_{:=\mathbf{C_n}}\mathbf{x_c}(nI) \\
&\underbrace{- \int_{(n-1)I-t_n}^{nI}\int_{t}^{nI-t_n} \mathbf{C_c} e^{\mathbf{A_c}(t-t')}\mathbf{B_c} d\mathbf{W_c}(t')dt + \int_{(n-1)I-t_n}^{nI} \mathbf{D_c} d\mathbf{V_c}(t)}_{:=\mathbf{v}[n]}
\end{aligned}$$

$\mathbf{y_o}[n]$ is transmitted over the erasure channel, and the intermittent system $(\mathbf{A_c}, \mathbf{B_c}, \mathbf{C_n})$ with nonuniform samples and erasure probability p_e has the following system equations which correspond to (2.26) and (2.27).

$$\begin{aligned}
d\mathbf{x_c}(t) &= \mathbf{A_c}\mathbf{x_c}(t)dt + \mathbf{B_c}d\mathbf{W_c}(t) \\
\mathbf{y}[n] &= \beta[n](\mathbf{C_n}\mathbf{x_c}(nI) + \mathbf{v}[n])
\end{aligned}$$

Then, the intermittent observability condition for $(\mathbf{A_c}, \mathbf{B_c}, \mathbf{C_n})$ is similar to Theorem 2.8.

Proposition 2.4. *Let t_n be i.i.d. random variables uniformly distributed on $[0, T]$ $(T > 0)$, and $(\mathbf{A_c}, \mathbf{B_c})$ be controllable. If $(\mathbf{A_c}, \mathbf{C_c})$ has unobservable and unstable eigenvalues, the intermittent system $(\mathbf{A_c}, \mathbf{B_c}, \mathbf{C_n})$ with nonuniform samples is not intermittent observable for all p_e. Otherwise, the intermittent system $(\mathbf{A_c}, \mathbf{B_c}, \mathbf{C_n})$ with nonuniform samples is intermittent observable if and only if $p_e < \frac{1}{|e^{2\lambda_{max}I}|}$ where λ_{max} is the eigenvalue of $\mathbf{A_c}$ with the largest real part.*

Compared to Theorem 2.8, we can see that the observability condition of $(\mathbf{A_c}, \mathbf{C})$ is relaxed to the observability condition of $(\mathbf{A_c}, \mathbf{C_c})$. This is due to the following fact: $\int_{(n-1)I-t_n}^{nI-t_n} e^{j\frac{2\pi}{T}t}dt = 0$ for all t_n and $\int_{(n-1)I-t_n}^{nI} e^{j\frac{2\pi}{T}t}dt \neq 0$ for some t_n. Even if $(\mathbf{A_c}, \mathbf{C_c})$ is observable, $(\mathbf{A_c}, \mathbf{C})$ can be unobservable for all t_n while $(\mathbf{A_c}, \mathbf{C_n})$ is observable for almost all t_n.

Nonuniform Time-varying Filtering

In some cases, it is impossible to change the sampling time. In this case, we can use nonuniform time-varying filtering to break eigenvalue cycles. Consider the following discrete-time system:

$$\begin{aligned}
\mathbf{x}[n+1] &= \mathbf{A}\mathbf{x}[n] + \mathbf{B}\mathbf{w}[n] \\
\mathbf{y_o}[n] &= \mathbf{C}\mathbf{x}[n] + \mathbf{v}[n]
\end{aligned}$$

Here $\mathbf{y_o}[n]$ are the observations at the sensor, and the sensor cannot change the sampling intervals. Instead, the sensor introduces nonuniform filtering to the observations as follows:

$$\mathbf{y'_o}[n] = \alpha[n]\mathbf{y_o}[n] + \alpha'[n]\mathbf{y_o}[n-1]$$

This is just like introducing an FIR (finite impulse response) filter at the sensor except that the impulse response of the filter keeps changing over time.

The output of the nonuniform time-varying filter, $\mathbf{y'_o}[n]$, is transmitted over the erasure channel. Therefore, the intermittent system $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ with erasure probability p_e and nonuniform time-varying filtering has the following system equations:

$$\begin{aligned}
&\mathbf{x}[n+1] = \mathbf{A}\mathbf{x}[n] + \mathbf{B}\mathbf{w}[n] \\
&\mathbf{y}[n] = \beta[n](\mathbf{y'_o}[n]) \\
&\quad = \beta[n](\alpha[n]\mathbf{C}\mathbf{x}[n] + \alpha'[n]\mathbf{C}\mathbf{x}[n-1] + \alpha[n]\mathbf{v}[n] + \alpha'[n]\mathbf{v}[n-1])
\end{aligned}$$

The intermittent observability with nonuniform filtering is given as the following proposition.

Proposition 2.5. *Let $\alpha[n]$ and $\alpha'[n]$ be i.i.d. random variables uniformly distributed on $[0, T]$ $(T > 0)$, and $(\mathbf{A}, \mathbf{B})$ be controllable. If $(\mathbf{A}, \mathbf{C})$ has unobservable and unstable eigenvalues, the intermittent system $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ with nonuniform filtering is not intermittent observable for all p_e. Otherwise, the intermittent system $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ with nonuniform filtering is intermittent observable if and only if $p_e < \frac{1}{|\lambda_{max}|^2}$ where λ_{max} is the largest magnitude eigenvalue of $\mathbf{A}$.*

Sampling with Nonuniform Waveforms

So far in Theorem 2.8, Proposition 2.4, and Proposition 2.5, we have seen three different ways of breaking eigenvalue cycles. However, these methods are essentially the same and generalized to nonuniform sampling with nonuniform waveforms.

Fig. 2.4 shows the nonuniform sampling methods used to break eigenvalue cycles with respect to their waveforms. First, Fig. 2.4a shows the uniform sampling which is implicitly used to make discrete-time system (2.1), (2.2) from the underlying continuous-time system. As we saw in Theorem 2.7, the eigenvalue cycles were not broken in this case. Fig. 2.4b shows the nonuniform sampling by changing the starting time of the integration, which is used in Theorem 2.8. In this case, the eigenvalue cycles were successfully broken, but we can still observe the regularity in the integration intervals. Due to this regularity, we needed the observability of $(\mathbf{A_c}, \mathbf{C})$ instead of the observability of $(\mathbf{A_c}, \mathbf{C_c})$. Fig. 2.4c shows the nonuniform sampling by changing the integration interval, which is used in Proposition 2.4. The eigenvalue cycles were also broken in this case and due to the lack of regularity in sampling intervals, the observability of $(\mathbf{A_c}, \mathbf{C_c})$ was enough. Fig. 2.4d shows nonuniform filtering, which is used in Proposition 2.5 and successfully breaks the eigenvalue cycles. Therefore, we can conclude that as long as the sampling waveforms are not uniform as

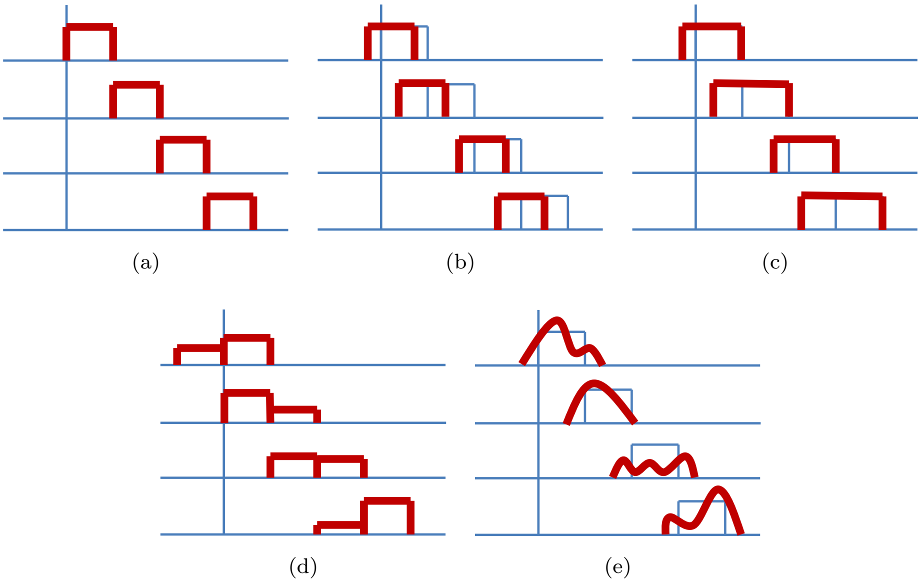


Figure 2.4: (a): uniform sampling of Theorem 2.7, (b): nonuniform sampling of Theorem 2.8, (c): nonuniform sampling of Proposition 2.4, (d): nonuniform filtering of Proposition 2.5, (e): nonuniform sampling with nonuniform waveforms

Fig. 2.4a the eigenvalue cycles are broken. In general, nonuniform waveforms shown in Fig. 2.4e can be used to break eigenvalue cycles, and it is an interesting technical equation to find the minimal condition on nonuniform waveforms to break eigenvalue cycles.

Extension to Parallel Channels

Theorem 2.8 can also be extended to the multiple sensors that transmit their observations through parallel erasure channels. Consider the following continuous-time system equations.

$$\begin{aligned}
d\mathbf{x_c}(t) &= \mathbf{A_c}\mathbf{x_c}(t)dt + \mathbf{B_c}d\mathbf{W_c}(t) \\
\mathbf{y_{c,1}}(t) &= \mathbf{C_{c,1}}\mathbf{x_c}(t) + \mathbf{D_{c,1}}\frac{d\mathbf{V_{c,1}(t)}}{dt} \\
&\vdots \\
\mathbf{y_{c,d}}(t) &= \mathbf{C_{c,d}}\mathbf{x_c}(t) + \mathbf{D_{c,d}}\frac{d\mathbf{V_{c,d}(t)}}{dt}
\end{aligned}$$

Here t is non-negative real-valued time index. $\mathbf{A_c} \in \mathbb{C}^{m\times m}$, $\mathbf{B_c} \in \mathbb{C}^{m\times g}$, $\mathbf{C_{c,i}} \in \mathbb{C}^{l_i\times m}$ and $\mathbf{D_{c,i}} \in \mathbb{C}^{l_i\times l_i}$ where $\mathbf{D_{c,i}}$ is invertible. $\mathbf{W_c}(t)$ and $\mathbf{V_{c,1}}(t)$ are independent g and l_i-dimensional standard Wiener process respectively.

Like (2.25), the nth sample at the sensor i is obtained by integrating $\mathbf{y_{c,i}}(t)$ from $(n-1)I-$

$t_{n,i}$ to $nI - t_{n,i}$:

$$\begin{aligned}
\mathbf{y_{o,i}}[n] &= \int_{(n-1)I-t_{n,i}}^{nI-t_{n,i}} \mathbf{y_{c,i}}(t)dt \\
&= \underbrace{\left(\int_0^I \mathbf{C_{c,i}} e^{\mathbf{A_c}(t-I)} dt\right)}_{:=\mathbf{C_i}} \mathbf{x_c}(nI - t_{n,i}) \\
&\quad \underbrace{- \int_{(n-1)I-t_{n,i}}^{nI-t_{n,i}} \int_t^{nI-t_{n,i}} \mathbf{C_{c,i}} e^{\mathbf{A_c}(t-t')} \mathbf{B_c} d\mathbf{W_c}(t')dt + \int_{(n-1)I-t_{n,i}}^{nI-t_{n,i}} \mathbf{D_{c,i}} d\mathbf{V_{c,i}}(t)}_{:=\mathbf{v_i}[n]}
\end{aligned}$$

Since $\mathbf{y_{o,i}}[n]$ are transmitted over the parallel erasure channel, the intermittent system $(\mathbf{A_c}, \mathbf{B_c}, \mathbf{C_i})$ with parallel channel has the following system equation:

$$\begin{aligned}
& d\mathbf{x_c}(t) = \mathbf{A_c}\mathbf{x_c}(t)dt + \mathbf{B_c}d\mathbf{W_c}(t) \\
& \mathbf{y_1}[n] = \beta_1[n](\mathbf{C_1}\mathbf{x_c}(nI - t_{n,1}) + \mathbf{v_1}[n]) \\
& \vdots \\
& \mathbf{y_d}[n] = \beta_d[n](\mathbf{C_d}\mathbf{x_c}(nI - t_{n,d}) + \mathbf{v_d}[n])
\end{aligned}$$

where $\mathbf{y_i}[n] \in \mathbb{C}^{l_i}$ and $\beta_i[n]$ are independent Bernoulli random processes with erasure probability $p_{e,i}$.

Like before, by a change of coordinates, we can rewrite the above system equations to the ones with a Jordan form $\mathbf{A_c}$ without changing the intermittent observability. Therefore, like (2.16), (2.17) and (2.18) we can write $\mathbf{A_c}$ and $\mathbf{C_i}$ as follows without loss of generality.

$$\begin{aligned}
& \mathbf{A_c} = diag\{\mathbf{A_{1,1}}, \mathbf{A_{1,2}}, \cdots, \mathbf{A}_{\mu,\nu_\mu}\} \\
& \mathbf{C_i} = \begin{bmatrix} \mathbf{C_{1,1,i}} & \mathbf{C_{1,2,i}} & \cdots & \mathbf{C}_{\mu,\nu_\mu,\mathbf{i}} \end{bmatrix}
\end{aligned}$$

where

$\mathbf{A_{i,j}}$ is a Jordan block with eigenvalue λ_i

$\lambda_1, \cdots, \lambda_\mu$ are pairwise distinct

$\mathbf{C_{i,j,k}}$ is a $l_k \times \dim \mathbf{A_{i,j}}$ complex matrix.

Denote

$$\mathbf{C_{i,j}} = \begin{bmatrix} (\mathbf{C_{i,1,j}})_1 & \cdots & (\mathbf{C}_{\mathbf{i},\nu_\mathbf{i},\mathbf{j}})_1 \end{bmatrix}$$

where $(\mathbf{C_{i,j,k}})_1$ implies the first column of $\mathbf{C_{i,j,k}}$

Let $(l_{i,1}, l_{i,2}, \cdots, l_{i,d}) \in \{0,1\}^d$ such that

$$\begin{bmatrix} \mathbf{1}(l_{i,1} = 0)\mathbf{C_{i,1}} \\ \vdots \\ \mathbf{1}(l_{i,d} = 0)\mathbf{C_{i,d}} \end{bmatrix}$$

is rank deficient, i.e. the rank is strictly less than ν_i.

Denote L_i as the set of such $(l_{i,1}, l_{i,2}, \cdots, l_{i,d})$ vectors. Then, the intermittent observability of the system $(\mathbf{A_c}, \mathbf{B_c}, \mathbf{C_i})$ with parallel channels is characterized by the following proposition.

Proposition 2.6. *Given an intermittent system* $(\mathbf{A_c}, \mathbf{B_c}, \mathbf{C_i})$ *with probability of erasures* $(p_{e,1}, \cdots, p_{e,d})$, *let* $(\mathbf{A_c}, \mathbf{B_c})$ *be controllable, and* $t_{n,i}$ *be independent random variables uniformly distributed on* $[0, T]$ $(T > 0)$. *The intermittent system* $(\mathbf{A_c}, \mathbf{B_c}, \mathbf{C_i})$ *with parallel channels is intermittent observable if and only if*

$$\max_{1 \le i \le \mu} \max_{(l_{i,1}, l_{i,2}, \cdots, l_{i,d}) \in L_i} \left(\prod_{1 \le j \le d} p_{e,j}^{l_{i,j}} \right) |e^{2\lambda_i I}| < 1.$$

2.8 Proofs

The proofs of Theorem 2.7 and Theorem 2.8 are quite similar. For presentation purposes, we will first present the proof of the nonuniform sampling case, Theorem 2.8, which is easier than that of Theorem 2.7. The randomness introduced by non-uniform sampling will be emulated in the proof of Theorem 2.7 by using Weyl's criterion [54].

2.8.1 Sufficiency Proof of Theorem 2.8 (Non-uniform Sampling)

We will prove that if $(\mathbf{A_c}, \mathbf{C})$ does not have unobservable and unstable eigenvalues and $p_e < \frac{1}{|e^{2\lambda_{max} I}|}$, the system is intermittent observable.

• Reduction to a Jordan form matrix $\mathbf{A_c}$: To simplify the problem, we first restrict to system equations (2.26) and (2.27) with the following properties. We will also justify that this restriction is without loss of generality and does not change intermittent observability.

(a) The system matrix $\mathbf{A_c}$ is a Jordan form matrix.

(b) All eigenvalues of $\mathbf{A_c}$ are unstable, i.e. the real parts are nonnegative.

(c) (2.26) and (2.27) can be extended to two-sided processes. (i.e. We can extend time to be negative, and set the state as zero there.)

The restriction (a) can be justified by a similarity transform [17]. As mentioned before, it is known [17] that for any square matrix $\mathbf{A_c}$, there exists an invertible matrix $\mathbf{U}$ and an upper-triangular Jordan matrix $\mathbf{A'_c}$ such that $\mathbf{A_c} = \mathbf{U}\mathbf{A'_c}\mathbf{U}^{-1}$. Then, equations (2.21) and (2.25) can be rewritten as

$$\begin{aligned}
\mathbf{U}^{-1}\mathbf{x_c}(t) &= e^{\mathbf{A'_c}t}\mathbf{U}^{-1}\mathbf{x_c}(0) + \int_0^t e^{\mathbf{A'_c}(t-t')}\mathbf{U}^{-1}\mathbf{B_c}d\mathbf{W_c}(t') \\
\mathbf{y_o}[n] &= \int_0^I \mathbf{C_c}\mathbf{U}e^{\mathbf{A'_c}(t-I)}dt\mathbf{U}^{-1}\mathbf{x_c}(nI - t_n) \\
&\quad - \int_{(n-1)I-t_n}^{nI-t_n} \int_t^{nI-t_n} \mathbf{C_c}\mathbf{U}e^{\mathbf{A'_c}(t-t')}\mathbf{U}^{-1}\mathbf{B_c}d\mathbf{W_c}(t')dt + \int_{(n-1)I-t_n}^{nI-t_n} \mathbf{D_c}d\mathbf{V_c}(t).
\end{aligned}$$

Thus, by denoting $\mathbf{x'_c}(t) := \mathbf{U}^{-1}\mathbf{x_c}(t)$, $\mathbf{B'_c} := \mathbf{U}^{-1}\mathbf{B_c}$, and $\mathbf{C'_c} := \mathbf{C_c}\mathbf{U}$, the system equations (2.19), (2.20) and (2.27) can be written in the following equivalent forms.

$$d\mathbf{x'_c}(t) = \mathbf{A'_c}\mathbf{x'_c}(t)dt + \mathbf{B'_c}d\mathbf{W_c}(t)$$
$$\mathbf{y_c}(t) = \mathbf{C'_c}\mathbf{x'_c}(t) + \mathbf{D_c}\frac{d\mathbf{V_c}(t)}{dt}$$
$$\mathbf{y_o}[n] = \mathbf{C}'\mathbf{x'_c}(nI - t_n) + \mathbf{v}[n]$$

where $\mathbf{C}' := \int_0^I \mathbf{C'_c} e^{\mathbf{A'_c}(t-I)}dt = \int_0^I \mathbf{C_c}\mathbf{U}\mathbf{U}^{-1}e^{\mathbf{A_c}(t-I)}\mathbf{U}dt = \mathbf{C}\mathbf{U}$.

Since $\mathbf{U}$ is invertible, $(\mathbf{A_c}, \mathbf{C})$ has an unobservable eigenvalue λ if and only if $(\mathbf{A'_c}, \mathbf{C}')$ has an unobservable eigenvalue λ. Moreover, since $\mathbf{x'_c} = \mathbf{U}^{-1}\mathbf{x_c}(t)$, the original intermittent system $(\mathbf{A_c}, \mathbf{B_c}, \mathbf{C})$ with nonuniform samples is intermittent observable if and only if the new intermittent system $(\mathbf{A'_c}, \mathbf{B'_c}, \mathbf{C}')$ with nonuniform samples is intermittent observable. Thus, without loss of generality, we can assume $\mathbf{A_c}$ is given in a Jordan form, which justifies (a).

Once $\mathbf{A_c}$ is given in a Jordan form, there is a natural correspondence between the eigenvalues and the states. If there is a stable eigenvalue — i.e. the real part of the eigenvalue is negative —, the variance of the corresponding state is uniformly bounded. Thus, we do not have to estimate such a state to make the estimation error finite. In the observation $\mathbf{y}[n]$, the stable states can be considered as a part of observation noise $\mathbf{v}[n]$, and the variance of $\mathbf{v}[n]$ is still uniformly bounded (even if $\mathbf{v}[n]$ can be correlated). Therefore, we can assume (b) without loss of generality.

To justify restriction (c), we set $\mathbf{W_c}(t) = 0$ for $t < 0$, $\mathbf{V_c}(t) = 0$ for $t < 0$, and let $\beta[n]$ be a two-sided Bernoulli process with probability $1 - p_e$. Then, the resulting two-sided processes $\mathbf{x_c}(t)$ and $\mathbf{y}[n]$ are identical to the original one-sided processes except that $\mathbf{x_c}(t) = 0$ for $t \in \mathbb{R}^{--}$ and $\mathbf{y}[n] = 0$ for $n \in \mathbb{Z}^{--}$.

In summary, without loss of generality we can assume that $\mathbf{A_c}$ is in a Jordan form, all eigenvalues of $\mathbf{A_c}$ are stable, and (2.26) and (2.27) are two-sided processes. Thus, we can assume $\mathbf{A_c} \in \mathbb{C}^{m\times m}$ and $\mathbf{C} \in \mathbb{C}^{l\times m}$ is given as follows.

$$\mathbf{A_c} = diag\{\mathbf{A_{1,1}}, \mathbf{A_{1,2}}, \cdots, \mathbf{A}_{\mathbf{1},\nu_\mathbf{1}}, \cdots, \mathbf{A}_{\mu,\mathbf{1}}, \cdots, \mathbf{A}_{\mu,\nu_\mu}\} \tag{2.28}$$
$$\mathbf{C} = \begin{bmatrix} \mathbf{C_{1,1}} & \mathbf{C_{1,2}} & \cdots & \mathbf{C}_{\mathbf{1},\nu_\mathbf{1}} & \cdots & \mathbf{C}_{\mu,\mathbf{1}} & \cdots & \mathbf{C}_{\mu,\nu_\mu} \end{bmatrix} \tag{2.29}$$

where

$\mathbf{A_{i,j}}$ is a Jordan block with eigenvalue $\lambda_i + j\omega_i$ and size $m_{i,j}$

$m_{i,1} \leq m_{i,2} \leq \cdots \leq m_{i,\nu_i}$ for all $i = 1, \cdots, \mu$

$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_\mu \geq 0$

$\lambda_1 + j\omega_1, \lambda_2 + j\omega_2, \cdots, \lambda_\mu + j\omega_\mu$ are pairwise distinct

$\mathbf{C_{i,j}}$ is a $l \times m_{i,j}$ complex matrix

The first columns of $\mathbf{C_{i,1}}, \mathbf{C_{i,2}}, \cdots, \mathbf{C}_{\mathbf{i},\nu_\mathbf{i}}$ are linearly independent.

Here, $\mathbf{A_{i,1}}, \cdots, \mathbf{A_{i,\nu_i}}$ are the Jordan blocks corresponding to the same eigenvalue. The Jordan blocks are sorted by the real parts of the eigenvalues in a descending order. The linear independence of $\mathbf{C_{i,1}}, \mathbf{C_{i,2}}, \cdots, \mathbf{C_{i,\nu_i}}$ comes from the observability of $(\mathbf{A_c}, \mathbf{C})$ (by Theorem 2.6).

• Uniform boundedness of observation noise: To prove intermittent observability, we will propose a suboptimal maximum-likelihood-style estimator, and analyze it. To upper bound the estimation error, we upper bound the disturbances and observation noises in the system.

By (2.22), we have

$$\mathbf{x_c}((n-k)I - t_{n-k}) = e^{-\mathbf{A_c}(kI+t_{n-k})}\mathbf{x_c}(nI) \underbrace{- \int_{(n-k)I-t_{n-k}}^{nI} e^{\mathbf{A_c}((n-k)I-t_{n-k}-t')}\mathbf{B_c}d\mathbf{W_c}(t')}_{:=\mathbf{w}'[n-k]}.$$

By plugging this equation into (2.27), we get

$$\begin{aligned} \mathbf{y}[n-k] &= \mathbf{C}\mathbf{x_c}((n-k)I - t_{n-k}) + \mathbf{v}[n-k] \\ &= \mathbf{C}e^{-\mathbf{A_c}(kI+t_{n-k})}\mathbf{x_c}(nI) + \underbrace{\mathbf{C}\mathbf{w}'[n-k] + \mathbf{v}[n-k]}_{:=\mathbf{v}'[n-k]}. \end{aligned} \tag{2.30}$$

We will upper bound the variance of $\mathbf{v}'[n-k]$. First, consider the variance of $w'[n-k]$. By assumption (b), all eigenvalues of $\mathbf{A_c}$ are unstable, and since $t_{n-k} \in [0,T]$, $((n-k)I - t_{n-k} - t')$ is within $[-(kI+T), 0]$. Thus, there exits $p' \in \mathbb{N}$ such that

$$\mathbb{E}[\mathbf{w}'[n-k]^\dagger \mathbf{w}'[n-k]] \lesssim 1 + k^{p'} \tag{2.31}$$

where $\lesssim$ holds for all n. (See Definition 2.12 for the definition of $\lesssim$.)

By (2.25), the variance of $\mathbf{v}[n]$ is uniformly bounded[14] for all n. Therefore, we have $\mathbb{E}[\mathbf{v}'[n-k]^\dagger \mathbf{v}'[n-k]] \lesssim 1 + k^{p'}$ for all n.

Moreover, since $W_c(t)$ is a standard Wiener process with unit variance, $\sup_{n\in\mathbb{Z}} \mathbb{E}[(\mathbf{x}(nI) - \widehat{\mathbf{x}}(nI))^\dagger(\mathbf{x}(nI) - \widehat{\mathbf{x}}(nI))] < \infty$ implies $\sup_{t\in\mathbb{R}} \mathbb{E}[(\mathbf{x}(t) - \widehat{\mathbf{x}}(t))^\dagger(\mathbf{x}(t) - \widehat{\mathbf{x}}(t))] < \infty$. Thus, it is enough to estimate the state only at discrete time steps.

• Suboptimal Maximum-Likelihood-Style Estimator: Now, we will give the suboptimal state estimator which only uses a finite number of recent observations. We first need the following key lemma.

Lemma 2.2. *Let $\mathbf{A_c}$ and $\mathbf{C}$ be given as in* (2.28) *and* (2.29)*, $\beta[n]$ be a Bernoulli process with probability $1-p_e$, and t_n be i.i.d. random variables whose distribution is uniform on $[0,T]$ $(T>0)$. Then, we can find $m' \in \mathbb{N}$, a polynomial $p(k)$ and a family of stopping times $\{S(\epsilon,k) : k \in \mathbb{Z}^+, 0 < \epsilon < 1\}$ such that for all $k \in \mathbb{Z}^+$ and $0 < \epsilon < 1$ there exist $k \le k_1 < k_2 < \cdots < k_{m'} \le S(\epsilon,k)$ and a $m \times m'l$ matrix $\mathbf{M}$ satisfying the following four conditions:*

(i) $\beta[k_i] = 1$ for all $1 \le i \le m'$

[14]Recall that to justify assumption (b), we considered the stable states as a part of observation noise $\mathbf{v}[n]$. However, this does not change the uniform boundedness since the variances of the stable states are also uniformly bounded.

(ii) $\mathbf{M}\begin{bmatrix}\mathbf{C}e^{-(k_1 I+t_{k_1})\mathbf{A_c}}\\ \mathbf{C}e^{-(k_2 I+t_{k_2})\mathbf{A_c}}\\ \vdots \\ \mathbf{C}e^{-(k_{m'} I+t_{k_{m'}})\mathbf{A_c}}\end{bmatrix}=\mathbf{I}_{m\times m}$

(iii) $|\mathbf{M}|_{max} \leq \frac{p(S(\epsilon,k))}{\epsilon} e^{\lambda_1 S(\epsilon,k)I}$

(iv) $\lim_{\epsilon \downarrow 0}\left(\exp \limsup_{s\to\infty} \sup_{k\in\mathbb{Z}^+} \frac{1}{s}\log \mathbb{P}\{S(\epsilon,k)-k=s\}\right) \leq p_e$.

Proof. See Appendix 7.4. The key ideas are explained in Section 2.5. □

Since we have $p_e < \frac{1}{|e^{2\lambda_{max}I}|} = \frac{1}{e^{2\lambda_1 I}}$, there exists $\delta > 1$ such that $\delta^5 p_e < \frac{1}{e^{2\lambda_1 I}}$. By Lemma 2.2, we can find $m' \in \mathbb{N}$, $0 < \epsilon < 1$, a polynomial $p(k)$ and a family of stopping times $\{S(n) : n \in \mathbb{Z}^+\}$ such that for all n, there exist $0 \leq k_1 < k_2 < \cdots < k_{m'} \leq S(n)$ and a $m \times m'l$ matrix $\mathbf{M_n}$ satisfying the following four conditions:

(i') $\beta[n-k_i]=1$ for $1 \leq i \leq m'$

(ii') $\mathbf{M_n}\begin{bmatrix}\mathbf{C}e^{-(k_1 I+t_{n-k_1})\mathbf{A_c}}\\ \mathbf{C}e^{-(k_2 I+t_{n-k_2})\mathbf{A_c}}\\ \vdots \\ \mathbf{C}e^{-(k_{m'} I+t_{n-k_{m'}})\mathbf{A_c}}\end{bmatrix}=\mathbf{I}_{m\times m}$

(iii') $|\mathbf{M_n}|_{max} \leq \frac{p(S(n))}{\epsilon} e^{\lambda_1 I\cdot S(n)}$

(iv') $\exp\left(\limsup_{s\to\infty} \sup_{n\in\mathbb{Z}^+} \frac{1}{s}\log \mathbb{P}\{S(n)=s\}\right) \leq \sqrt{\delta} p_e$.

Then, here is the proposed suboptimal maximum likelihood estimator for $\mathbf{x}(nI)$:

$$\widehat{\mathbf{x}}(nI)=\mathbf{M_n}\begin{bmatrix}\mathbf{y}[n-k_1]\\ \mathbf{y}[n-k_2]\\ \vdots \\ \mathbf{y}[n-k_{m'}]\end{bmatrix}. \tag{2.32}$$

Here, k_i also depends on n, but we omit the dependency in notation for simplicity. Notice that, m', the number of observations used is much larger than the dimension of the system, m. In other words, the estimator proposed here may use many more observations than the number of states (the number of observations that a simple matrix inverse observer needs). This is because we use a successive decoding idea in the proof of Lemma 2.2.

• Analysis of the estimation error: Now, we will analyze the performance of the proposed estimator. Recall that p' is defined in (2.31) and $\delta > 1$. By (iv') and well-known properties of polynomial and exponential functions, we can find $c > 0$ that satisfies the following three conditions:

(i") $(1+k^{p'}) \leq c\cdot\delta^k$ for all $k \geq 0$

(ii") $p(k) \leq c\cdot\delta^k$ for all $k \geq 0$

(iii") $\sup_{n\in\mathbb{N}} \mathbb{P}\{S(n)=s\} \leq c\cdot(\delta\cdot p_e)^s$ for all $s \in \mathbb{Z}^+$

Let $\mathcal{F}_\beta$ be the σ-field generated by $\beta[n]$ and t_i. Then, k_i, $S(n)$, and t_i are deterministic variables conditioned on $\mathcal{F}_\beta$. The estimation error is upper bounded by

$$
\begin{aligned}
&\sup_n \mathbb{E}[|\mathbf{x}(nI) - \widehat{\mathbf{x}}(nI)|_2^2] \\
&= \sup_n \mathbb{E}[\mathbb{E}[|\mathbf{x}(nI) - \widehat{\mathbf{x}}(nI)|_2^2 | \mathcal{F}_\beta]] \\
&\overset{(A)}{=} \sup_n \mathbb{E}[\mathbb{E}[\left| \mathbf{x}(nI) - \mathbf{M_n}(\begin{bmatrix} \mathbf{C}e^{-\mathbf{A_c}(k_1 I + t_{n-k_1})} \\ \mathbf{C}e^{-\mathbf{A_c}(k_2 I + t_{n-k_2})} \\ \vdots \\ \mathbf{C}e^{-\mathbf{A_c}(k_{m'} I + t_{n-k_{m'}})} \end{bmatrix} \mathbf{x}(nI) + \begin{bmatrix} \mathbf{v}'[n-k_1] \\ \mathbf{v}'[n-k_2] \\ \vdots \\ \mathbf{v}'[n-k_{m'}] \end{bmatrix}) \right|_2^2 | \mathcal{F}_\beta]] \\
&\overset{(B)}{=} \sup_n \mathbb{E}[\mathbb{E}[\left| \mathbf{M_n} \begin{bmatrix} \mathbf{v}'[n-k_1] \\ \mathbf{v}'[n-k_2] \\ \vdots \\ \mathbf{v}'[n-k_{\sum_{1 \le i \le \mu} m_i'}] \end{bmatrix} \right|_2^2 | \mathcal{F}_\beta]] \\
&\lesssim \sup_n \mathbb{E}[|\mathbf{M_n}|_{max}^2 \cdot \mathbb{E}[\left| \begin{bmatrix} \mathbf{v}'[n-k_1] \\ \mathbf{v}'[n-k_2] \\ \vdots \\ \mathbf{v}'[n-k_{m'}] \end{bmatrix} \right|_{max}^2 | \mathcal{F}_\beta]] \\
&\overset{(C)}{\lesssim} \sup_n \mathbb{E}[|\mathbf{M_n}|_{max}^2 \cdot (1 + S(n)^{p'})^2] \\
&\overset{(D)}{\le} \sup_n \mathbb{E}[\left(\frac{p(S(n))}{\epsilon} e^{\lambda_1 I \cdot S(n)} \right)^2 \cdot (1 + S(n)^{p'})^2] \\
&\overset{(E)}{\lesssim} \sup_n \mathbb{E}[\delta^{2S(n)} \cdot e^{2\lambda_1 I \cdot S(n)} \cdot \delta^{2S(n)}] \\
&\overset{(F)}{\lesssim} \sum_{s=0}^{\infty} \delta^{4s} \cdot e^{2\lambda_1 I \cdot s} \cdot (\delta \cdot p_e)^s \\
&\overset{(G)}{=} \sum_{s=0}^{\infty} (\delta^5 \cdot e^{2\lambda_1 I} \cdot p_e)^s \\
&< \infty
\end{aligned}
$$

where $\lesssim$ holds for all n.

(A): By (2.30) and (2.32).

(B): By condition (ii').

(C): Since $\mathbb{E}[\mathbf{v}'[n-k]^\dagger \mathbf{v}'[n-k]] \lesssim 1 + k^{p'}$ by definition.

(D): By condition (iii').

(E): By condition (i") and (ii").

(F): By condition (iii").

(G): Since we chose δ so that $\delta^5 p_e \cdot e^{2\lambda_1 I} < 1$.

Therefore, the estimation error is uniformly bounded over $t \in \mathbb{R}^+$ when $p_e < \frac{1}{e^{2\lambda_1 I}}$, which finishes the proof.

2.8.2 Necessity Proof of Theorem 2.8

The necessity proof divides into two parts. First, we prove that if $p_e \geq \frac{1}{|e^{2\lambda_{max} I}|}$, then the system is not intermittent observable. Second, we prove that if $(\mathbf{A_c}, \mathbf{C})$ has unobservable and unstable eigenvalues — i.e. $\exists \lambda \in \mathbb{C}^+$ such that $\begin{bmatrix} \lambda \mathbf{I} - \mathbf{A_c} \\ \mathbf{C} \end{bmatrix}$ is rank deficient — then the system is not intermittent observable.

• When $p_e \geq \frac{1}{|e^{2\lambda_{max} I}|}$: Intuitively speaking, we will give all states except the one corresponding to the maximum eigenvalue as side-information to the estimator. Thus, we will reduce the problem to the scalar system discussed in Section 2.5.

Formally, let $\mathbf{\Sigma_{t|t}} := \mathbb{E}[(\mathbf{x_c}(t) - \mathbb{E}[\mathbf{x_c}(t)|\mathbf{y}^{\lfloor \frac{t}{I} \rfloor}])(\mathbf{x_c}(t) - \mathbb{E}[\mathbf{x_c}(t)|\mathbf{y}^{\lfloor \frac{t}{I} \rfloor}])^\dagger | \mathcal{F}_\beta]$ where $\mathcal{F}_\beta$ is the σ-field generated by $\beta[n]$ and t_i. Notice that $\mathbf{\Sigma_{t|t}}$ is a random variable.

It is known that when $(\mathbf{A_c}, \mathbf{B_c})$ is controllable, the estimation error covariance of $\mathbf{x_c}(t)$ based on all the causally available information $\mathbf{y_c}(0:t)$ is positive definite when t is large enough. Therefore, there exists $t' > 0$ and $\sigma^2 > 0$ such that for all $t \geq t'$, $\mathbf{\Sigma_{t|t}} \succeq \sigma^2 \mathbf{I}$ with probability one. Let $\mathbf{e}$ be a right eigenvector of $\mathbf{A_c}$ associated with the eigenvalue λ_{max}, i.e. $\mathbf{A_c e} = \lambda_{max} \mathbf{e}$. Then, we can find $\sigma'^2 > 0$ such that for all $t \geq t'$, $\mathbf{\Sigma_{t|t}} \succeq \sigma'^2 \mathbf{e}\mathbf{e}^\dagger$ with probability one.

Define the stopping time $S'_n := \inf\{k \in \mathbb{Z}^+ | \beta[n-k] = 1\}$ as the time until the most recent observation.

The observations between discrete time $n - S'_n + 1$ and n are all erased. This implies the estimation error is exponentially amplified by the system dynamics during this period. Thus, conditioned on $(n - S'_n)I \geq t'$, $\mathbf{\Sigma_{nI|nI}}$ is lower bounded as follows with probability one.[15]

$$\begin{aligned} \mathbb{E}[\mathbf{\Sigma_{nI|nI}} | S'_n, (n - S'_n)I \geq t'] &\succeq (e^{\mathbf{A_c}(S'_n I)}) \mathbf{\Sigma_{(n-S'_n)I|(n-S'_n)I}} (e^{\mathbf{A_c}(S'_n I)})^\dagger \\ &\succeq \sigma'^2 (e^{\mathbf{A_c}(S'_n I)}) \mathbf{e}\mathbf{e}^\dagger (e^{\mathbf{A_c}(S'_n I)})^\dagger \\ &\succeq \sigma'^2 |e^{2\lambda_{max} I}|^{S'_n} \mathbf{e}\mathbf{e}^\dagger \end{aligned}$$

Here we use the fact that when $\mathbf{e}$ is an eigenvector of $\mathbf{A_c}$ associated with an eigenvalue λ_{max}, $\mathbf{e}$ is also an eigenvector of $e^{\mathbf{A_c} t}$ associated with the eigenvalue $e^{\lambda_{max} t}$ for all t.

Since $p_e \geq \frac{1}{|e^{2\lambda_{max} I}|}$, the average estimation error is lower bounded as follows:

$$\begin{aligned} &\mathbb{E}[(\mathbf{x_c}(nI) - \mathbb{E}[\mathbf{x_c}(nI)|\mathbf{y}^n])^\dagger (\mathbf{x_c}(nI) - \mathbb{E}[\mathbf{x_c}(nI)|\mathbf{y}^n])] \\ &\geq \mathbb{E}[\sigma'^2 |e^{2\lambda_{max} I}|^{S'_n} |\mathbf{e}|^2 \cdot \mathbf{1}((n - S'_n)I \geq t')] \\ &\geq \sigma'^2 |\mathbf{e}|^2 \cdot \sum_{0 \leq s \leq \lfloor n - \frac{t}{I} \rfloor} |e^{2\lambda_{max} I}|^s \cdot (1 - p_e) p_e^s \\ &\geq \sigma'^2 |\mathbf{e}|^2 \cdot (1 - p_e) \cdot (\lfloor n - \frac{t}{I} \rfloor + 1) \end{aligned}$$

[15] The lower bound does not hold for $\Re(\lambda) = 0$ which induces $p_e = 1$. However, in this case we do not have any observations, so trivially the system is unstable.

Thus, the estimation error goes to infinity as $n \to \infty$, so the system is not intermittently observable.

• When $(\mathbf{A_c}, \mathbf{C})$ has unobservable and unstable eigenvalues: Now, we prove that if $(\mathbf{A_c}, \mathbf{C})$ has unobservable and unstable eigenvalues, the system is not intermittent observable. This seems trivial, but the original continuous-time system $(\mathbf{A_c}, \mathbf{C_c})$ can still be observable while the sampled system $(\mathbf{A_c}, \mathbf{C})$ is not. Thus, it still needs justification.

Let $\lambda \in \mathbb{C}^+$ be an unobservable and unstable eigenvalue. Then, $\begin{bmatrix} \lambda\mathbf{I} - \mathbf{A_c} \\ \mathbf{C} \end{bmatrix}$ is rank deficient, and we can find a nonzero vector $\mathbf{i}$ such that $\begin{bmatrix} \lambda\mathbf{I} - \mathbf{A_c} \\ \mathbf{C} \end{bmatrix} \mathbf{i} = \mathbf{0}$. Then, $\mathbf{i}$ satisfies $\mathbf{Ci} = \mathbf{0}$, $\mathbf{A_c i} = \lambda \mathbf{i}$, and we can notice that $\mathbf{C}e^{\mathbf{A_c}t}\mathbf{i} = e^{\lambda t}\mathbf{Ci} = \mathbf{0}$. We will prove that the uncertainty in the direction $\mathbf{i}$ is not observable by any observations.

By the controllability of $(\mathbf{A_c}, \mathbf{B_c})$, as above there exists t' such that for all $t \geq t'$, $\mathbf{x_c}(t) - \mathbb{E}[\mathbf{x_c}(t)|\mathbf{y_c}(0 : t)]$ has a positive definite covariance matrix. Therefore, we can write $\mathbf{x_c}(t) - \mathbb{E}[\mathbf{x_c}(t)|\mathbf{y_c}(0:t)] = \mathbf{i} \cdot x_c'(t) + \mathbf{x_c''}(t)$ where $x_c'(t)$, $\mathbf{x_c''}(t)$ and $\mathbf{y_c}(0:t)$ are independent and $\mathbb{E}[|x_c'(t)|^2] \geq \sigma''^2$ for some $\sigma''^2 > 0$ and all $t \geq t'$.

Then, we will prove that the sampled observations are independent from $x_c'(t)$. By (2.21) and (2.25), for all $\tau \leq (n-1)I - t_n$ we have

$$
\begin{aligned}
\mathbf{y_o}[n] &= \mathbf{C}(e^{\mathbf{A_c}(nI-t_n-\tau)}(\mathbf{x_c}(\tau) + \int_{\tau}^{nI-t_n} e^{\mathbf{A_c}(\tau-t')}\mathbf{B_c}d\mathbf{W_c}(t'))) \\
&- \int_{(n-1)I-t_n}^{nI-t_n}\int_{t}^{nI-t_n} \mathbf{C_c}e^{\mathbf{A_c}(t-t')}\mathbf{B_c}d\mathbf{W_c}(t')dt + \int_{(n-1)I-t_n}^{nI-t_n} \mathbf{D_c}d\mathbf{V_c}(t) \\
&= \mathbf{C}(e^{\mathbf{A_c}(nI-t_n-\tau)}(\mathbf{i}\cdot x_c'(\tau) + \mathbf{x_c''}(\tau) + \mathbf{E}[\mathbf{x_c}(\tau)|\mathbf{y_c}(0:\tau)] + \int_{\tau}^{nI-t_n} e^{\mathbf{A_c}(\tau-t')}\mathbf{B_c}d\mathbf{W_c}(t'))) \\
&- \int_{(n-1)I-t_n}^{nI-t_n}\int_{t}^{nI-t_n} \mathbf{C_c}e^{\mathbf{A_c}(t-t')}\mathbf{B_c}d\mathbf{W_c}(t')dt + \int_{(n-1)I-t_n}^{nI-t_n} \mathbf{D_c}d\mathbf{V_c}(t) \\
&= \mathbf{C}(e^{\mathbf{A_c}(nI-t_n-\tau)}(\mathbf{x_c''}(\tau) + \mathbf{E}[\mathbf{x_c}(\tau)|\mathbf{y_c}(0:\tau)] + \int_{\tau}^{nI-t_n} e^{\mathbf{A_c}(\tau-t')}\mathbf{B_c}d\mathbf{W_c}(t'))) \\
&- \int_{(n-1)I-t_n}^{nI-t_n}\int_{t}^{nI-t_n} \mathbf{C_c}e^{\mathbf{A_c}(t-t')}\mathbf{B_c}d\mathbf{W_c}(t')dt + \int_{(n-1)I-t_n}^{nI-t_n} \mathbf{D_c}d\mathbf{V_c}(t) \qquad (2.33)
\end{aligned}
$$

where the last equality comes from $\mathbf{C}e^{\mathbf{A_c}t}\mathbf{i} = 0$. Moreover, by causality and definitions, the last equation is independent of $x_c'(\tau)$.

Now, we will prove that the uncertainty $x_c'(\tau)$ can be arbitrarily amplified. Since t_i are uniform random variables on $[0, T]$, there exists a positive probability such that $(n-1)I - t_n \leq (n+n'-1)I - t_{n+n'}$ for all $n' \in \mathbb{N}$. Denote such an event as E. Then, by choosing n large enough so that $(n-1)I - t_n \geq t'$, we have the following lower bound on the estimation error for all

$t \geq (n-1)I - t_n$:

$$\begin{aligned}
&\mathbb{E}[|\mathbf{x_c}(t) - \mathbb{E}[\mathbf{x_c}(t)|\mathbf{y}^{\lfloor \frac{t}{I} \rfloor}]|^2] \\
&\geq \mathbb{E}[|\mathbf{x_c}(t) - \mathbb{E}[\mathbf{x_c}(t)|\mathbf{y}^{\lfloor \frac{t}{I} \rfloor}]|^2|E]\mathbb{P}(E) \\
&\overset{(a)}{\geq} \mathbb{E}[|e^{\mathbf{A_c}(t-((n-1)I-t_n))}\mathbf{i} \cdot x_c''((n-1)I-t_n)|^2|E]\mathbb{P}(E) \\
&= |e^{\lambda(t-((n-1)I-T))} \cdot \mathbf{i}|^2 \sigma''^2 \cdot \mathbb{P}(E)
\end{aligned} \tag{2.34}$$

(a): By (2.21), $\mathbf{x_c}(t) = e^{\mathbf{A_c}(t-((n-1)I-t_n))}\mathbf{x_c}((n-1)I-t_n) + \int_{(n-1)I-t_n}^{t} e^{\mathbf{A_c}((n-1)I-t_n-t')}\mathbf{B_c} d\mathbf{W_c}(t')$. Moreover, by definition, $x_c''((n-1)I-t_n)$ is independent from $\mathbf{y_c}(0:(n-1)I-t_n)$. By (2.33), $x_c''((n-1)I-t_n)$ is also independent from $\mathbf{y_o}[n], \mathbf{y_o}[n+1], \cdots$.

Since we can choose t arbitrarily large, this finishes the proof for $\Re(\lambda) > 0$. To prove for the case of $\Re(\lambda) = 0$, we can bound (2.34) more carefully and justify that independent estimation errors accumulate in the direction of $\mathbf{i}$. We omit the proof here since the argument is essentially equivalent to that of the well-known fact that an eigenvalue with zero real part is unstable in continuous-time systems.

2.8.3 Sufficiency Proof of Theorem 2.7 (Discrete-Time Systems)

We will prove that if $p_e < \frac{1}{\max_{1 \leq i \leq \mu} |\lambda_{i,1}|^{2\frac{p_i}{l_i}}}$ then the system is intermittent observable.

• Reduction to a Jordan form matrix $\mathbf{A}$: As in Section 2.8.1, we will restrict attention to system equations (2.1) and (2.2) with the following properties, and justify that such a restriction is without loss of generality and does not change the intermittent observability.
(a) The system matrix $\mathbf{A}$ is a Jordan form matrix.
(b) All eigenvalues of $\mathbf{A}$ are unstable, i.e. the magnitude of all eigenvalues are greater or equal to 1.
(c) (2.1) and (2.2) can be extended to two-sided processes.

The restriction (a) can be justified by a similarity transform [17]. It is known [17] that for any square matrix $\mathbf{A}$, there exists an invertible matrix $\mathbf{U}$ and an upper-triangular Jordan matrix $\mathbf{A'}$ such that $\mathbf{A} = \mathbf{U}\mathbf{A'}\mathbf{U}^{-1}$. Then, the system equations (2.1) and (2.2) can be rewritten as:

$$\begin{aligned}
&\mathbf{U}^{-1}\mathbf{x}[n+1] = \mathbf{A'}\mathbf{U}^{-1}\mathbf{x}[n] + \mathbf{U}^{-1}\mathbf{B}\mathbf{w}[n] \\
&\mathbf{y}[n] = \beta[n](\mathbf{C}\mathbf{U}\mathbf{U}^{-1}\mathbf{x}[n] + \mathbf{v}[n]).
\end{aligned}$$

Thus, by denoting $\mathbf{x'}[n] := \mathbf{U}^{-1}\mathbf{x}[n]$, $\mathbf{B'} := \mathbf{U}^{-1}\mathbf{B}$, and $\mathbf{C'} := \mathbf{C}\mathbf{U}$, we get

$$\begin{aligned}
&\mathbf{x'}[n+1] = \mathbf{A'}\mathbf{x'}[n] + \mathbf{B'}\mathbf{w}[n] \\
&\mathbf{y}[n] = \beta[n](\mathbf{C'}\mathbf{x'}[n] + \mathbf{v}[n]).
\end{aligned}$$

Since $\mathbf{U}$ is invertible, the controllability of $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ remains the same for the new intermittent system $(\mathbf{A'}, \mathbf{B'}, \mathbf{C'})$. Moreover, since $\mathbf{x'}[n] = \mathbf{U}^{-1}\mathbf{x}[n]$, the original intermittent system is intermittent observable if and only if the new intermittent system is intermittent observable. Thus, without loss of generality, we can assume that $\mathbf{A}$ is given in a Jordan form, which justifies (a).

Once $\mathbf{A}$ is given in Jordan form, there is a natural correspondence between the eigenvalues and the states. If there is a stable eigenvalue — i.e. the magnitude of the eigenvalue is less than 1 —, the variance of the corresponding state is uniformly bounded. Thus, we do not have to estimate that particular state to make the estimation error finite. In the observation $\mathbf{y}[n]$, the stable states can be considered as a part of observation noise $\mathbf{v}[n]$, and the variance of $\mathbf{v}[n]$ is still uniformly bounded (even if $\mathbf{v}[n]$ can be correlated). Therefore, we can assume (b) without loss of generality.

To justify restriction (c), rewrite (2.1) as

$$\mathbf{x}[n+1] = \mathbf{A}\mathbf{x}[n] + \mathbf{I}\mathbf{w}'[n]$$

where $\mathbf{w}'[n] = \mathbf{B}\mathbf{w}[n]$ for $n \geq 0$. Let $\mathbf{w}'[-1] = \mathbf{x}[0]$, $\mathbf{w}[n] = \mathbf{0}$ for $n < -1$, and $\mathbf{v}[n]$ for $n < 0$. We also extend $\beta[n]$ to a two-sided Bernoulli process with probability $1 - p_e$. Then, the resulting two-sided processes $\mathbf{x}[n]$ and $\mathbf{y}[n]$ are identical to the original one-sided processes except that $\mathbf{x}[n] = \mathbf{0}$ and $\mathbf{y}[n] = \mathbf{0}$ for $n \in \mathbb{Z}^{--}$.

In summary, without loss of generality we can assume that $\mathbf{A}$ is in a Jordan form, all eigenvalues of $\mathbf{A}$ is stable, and (2.1) and (2.2) are two-sided process. Therefore, we can assume that $\mathbf{A} \in \mathbb{C}^{m \times m}$ and $\mathbf{C} \in \mathbb{C}^{l \times m}$ are given as

$$\begin{aligned}
&\mathbf{A} = diag\{\mathbf{A_{1,1}}, \mathbf{A_{1,2}}, \cdots, \mathbf{A_{1,\nu_1}}, \cdots, \mathbf{A_{\mu,1}}, \cdots, \mathbf{A_{\mu,\nu_\mu}}\} \\
&\mathbf{C} = \begin{bmatrix} \mathbf{C_{1,1}} & \mathbf{C_{1,2}} & \cdots & \mathbf{C_{1,\nu_1}} & \cdots & \mathbf{C_{\mu,1}} & \cdots & \mathbf{C_{\mu,\nu_\mu}} \end{bmatrix} \\
&\text{where} \\
&\quad \mathbf{A_{i,j}} \text{ is a Jordan block with an eigenvalue } \lambda_{i,j} \text{ and size } m_{i,j} \\
&\quad m_{i,1} \geq m_{i,2} \geq \cdots \geq m_{i,\nu_i} \text{ for all } i = 1, \cdots, \mu \\
&\quad |\lambda_{1,1}| \geq |\lambda_{2,1}| \geq \cdots \geq |\lambda_{\mu,1}| \geq 1 \\
&\quad \{\lambda_{i,1}, \cdots, \lambda_{i,\nu_i}\} \text{ is cycle with length } \nu_i \text{ and period } p_i \\
&\quad \text{For } i \neq i',\ \{\lambda_{i,j}, \lambda_{i',j'}\} \text{ is not a cycle} \\
&\quad \mathbf{C_{i,j}} \text{ is a } l \times m_{i,j} \text{ complex matrix.}
\end{aligned} \tag{2.35}$$

Here, $\mathbf{A_{i,1}}, \cdots, \mathbf{A_{i,\nu_i}}$ are the Jordan blocks corresponding to the same eigenvalue cycle. The Jordan blocks are sorted in descending order by the magnitude of the eigenvalues.

Like (2.17), (2.18), we also define $\mathbf{A_i}$, $\mathbf{C_i}$, and l_i as follows.

$$\begin{aligned}
&\mathbf{A_i} = diag\{\lambda_{i,1}, \cdots, \lambda_{i,\nu_i}\} \\
&\mathbf{C_i} = \begin{bmatrix} (\mathbf{C_{i,1}})_1 & \cdots & (\mathbf{C_{i,\nu_i}})_1 \end{bmatrix} \\
&\text{where } (\mathbf{C_{i,j}})_1 \text{ is the first column of } \mathbf{C_{i,j}}.
\end{aligned} \tag{2.36}$$

l_i is the minimum cardinality among the sets $S' \subseteq \{0, 1, \cdots, p_i - 1\}$ whose resulting $S := \{0, 1, \cdots, p_i - 1\} \setminus S' = \{s_1, s_2, \cdots, s_{|S|}\}$ makes

$$\begin{bmatrix} \mathbf{C_i A_i}^{s_1} \\ \mathbf{C_i A_i}^{s_2} \\ \vdots \\ \mathbf{C_i A_i}^{s_{|S|}} \end{bmatrix} \tag{2.37}$$

be rank deficient, i.e. the rank of the matrix (2.37) is strictly less than ν_i.

Moreover, in (2.3), we already assumed that there exists a finite $\sigma > 0$ such that

$$\begin{aligned} &\sup_{n \in \mathbb{Z}} \mathbb{E}[\mathbf{w}[n]\mathbf{w}[n]^\dagger] \preceq \sigma^2 \mathbf{I} \\ &\sup_{n \in \mathbb{Z}} \mathbb{E}[\mathbf{v}[n]\mathbf{v}[n]^\dagger] \preceq \sigma^2 \mathbf{I}. \end{aligned} \tag{2.38}$$

• Uniform boundedness of observation noise: To prove intermittent observability, we will propose a suboptimal maximum-likelihood-style estimator, and analyze it. We first have to upper bound the disturbances and observation noises in the system. Following the same steps of (2.15), we can derive

$$\mathbf{y}[n-k] = \mathbf{C}\mathbf{A}^{-k}\mathbf{x}[n] - \underbrace{(\mathbf{C}\mathbf{A}^{-1}\mathbf{w}[n-k] + \cdots + \mathbf{C}\mathbf{A}^{-k}\mathbf{w}[n-1] - \mathbf{v}[n-k])}_{\mathbf{v}'[n-k]}. \tag{2.39}$$

The invertibility of $\mathbf{A}$ is comes from assumption (b). Moreover, since all eigenvalues of $\mathbf{A}$ are unstable, by (2.38) we can find $p' \in \mathbb{N}$ such that

$$\mathbb{E}[\mathbf{v}'[n-k]^\dagger \mathbf{v}'[n-k]] \lesssim 1 + k^{p'} \tag{2.40}$$

where $\lesssim$ holds for all $n, k (k \le n)$.

• Suboptimal Maximum-Likelihood-Style Estimator: Now, we will give a suboptimal estimator for the state which only uses a finite number of recent observations. We first need the following key lemma which plays a parallel role to Lemma 2.2.

Lemma 2.3. *Let* $\mathbf{A}$ *and* $\mathbf{C}$ *be given as in* (2.35), (2.36) *and* (2.37), *and* $\beta[n]$ *be a Bernoulli process with probability* $1 - p_e$. *Then, we can find* $m'_1, \cdots, m'_\mu \in \mathbb{N}$, *polynomials* $p_1(k), \cdots, p_\mu(k)$ *and families of stopping times* $\{S_1(\epsilon, k) : k \in \mathbb{Z}^+, 0 < \epsilon < 1\}, \cdots, \{S_\mu(\epsilon, k) : k \in \mathbb{Z}^+, 0 < \epsilon < 1\}$ *such that for all* $k \in \mathbb{Z}^+$ *and* $0 < \epsilon < 1$ *there exist* $k \le k_1 < \cdots < k_{m'_1} \le S_1(\epsilon, k) < k_{m'_1+1} < \cdots < k_{\sum_{1 \le i \le \mu} m'_i} \le S_\mu(\epsilon, k)$ *and an* $m \times (\sum_{1 \le i \le \mu} m'_i) l$ *matrix* $\mathbf{M}$ *satisfying the following conditions:*

(i) $\beta[k_i] = 1$ *for* $1 \le i \le \sum_{1 \le i \le \mu} m'_i$

(ii) $\mathbf{M} \begin{bmatrix} \mathbf{C}\mathbf{A}^{-k_1} \\ \mathbf{C}\mathbf{A}^{-k_2} \\ \vdots \\ \mathbf{C}\mathbf{A}^{-k_{\sum_{1 \le i \le \mu} m'_i}} \end{bmatrix} = \mathbf{I}_{m \times m}$

(iii) $|\mathbf{M}|_{max} \leq \max_{1 \leq i \leq \mu} \left\{ \frac{p_i(S_i(\epsilon,k))}{\epsilon} |\lambda_{i,1}|^{S_i(\epsilon,k)} \right\}$

(iv) $\lim_{\epsilon \downarrow 0} \exp\left(\limsup_{s \to \infty} \sup_{k \in \mathbb{Z}^+} \frac{1}{s} \log \mathbb{P}\{S_i(\epsilon,k) - k = s\}\right) \leq \max_{1 \leq j \leq i} \left\{ p_e^{\frac{l_j}{p_j}} \right\}$ *for* $1 \leq i \leq \mu$

(v) $\lim_{\epsilon \downarrow 0} \exp\left(\limsup_{s \to \infty} \operatorname{ess\,sup} \frac{1}{s} \log \mathbb{P}\{S_a(\epsilon,k) - S_b(\epsilon,k) = s | \mathcal{F}_{S_b}\}\right) \leq \max_{b < i \leq a} \left\{ p_e^{\frac{l_i}{p_i}} \right\}$ *for* $1 \leq b < a \leq \mu$ *where* $\mathcal{F}_{S_i}$ *is the* σ*-field generated by* $S_i(\epsilon,k)$*.*

Proof. See Appendix 7.7. The ideas in the proof are discussed in Section 2.5. □

Since $p_e < \frac{1}{\max_{1 \leq i \leq \mu} |\lambda_{i,1}|^{2\frac{p_i}{l_i}}}$, there exists $\delta > 1$ such that $\delta^5 \cdot \max_{1 \leq i \leq \mu} p_e^{\frac{l_i}{p_i}} |\lambda_{i,1}|^2 < 1$. By Lemma 2.3, we can find $m'_1, \cdots, m'_\mu \in \mathbb{N}$, $0 < \epsilon < 1$, polynomials $p_1(k), \cdots, p_\mu(k)$, and a family of stopping times $\{(S_1(n), \cdots, S_\mu(n)) : n \in \mathbb{Z}^+\}$ such that $\forall n$ there exist $0 \leq k_1 < \cdots < k_{m'_1} \leq S_1(n) < k_{m'_1+1} < \cdots < k_{\sum_{1 \leq i \leq \mu} m'_i} \leq S_\mu(n)$ and a $m \times (\sum_{1 \leq i \leq \mu} m'_i) l$ matrix $\mathbf{M_n}$ satisfying the following conditions:

(i') $\beta[n - k_i] = 1$ for $1 \leq i \leq \sum_{1 \leq i \leq \mu} m'_i$

(ii') $\mathbf{M_n} \begin{bmatrix} \mathbf{CA}^{-k_1} \\ \mathbf{CA}^{-k_2} \\ \vdots \\ \mathbf{CA}^{-k_{\sum_{1 \leq i \leq \mu} m'_i}} \end{bmatrix} = \mathbf{I}_{m \times m}$

(iii') $|\mathbf{M_n}|_{max} \leq \max_{1 \leq i \leq \mu} \left\{ \frac{p_i(S_i(n))}{\epsilon} |\lambda_{i,1}|^{S_i(n)} \right\}$

(iv') $\exp\left(\limsup_{s \to \infty} \frac{1}{s} \log \mathbb{P}\{S_i(n) = s\}\right) \leq \sqrt{\delta} \cdot \max_{1 \leq j \leq i} \left\{ p_e^{\frac{l_j}{p_j}} \right\}$ for $1 \leq i \leq \mu$

(v') $\exp\left(\limsup_{s \to \infty} \operatorname{ess\,sup} \frac{1}{s} \log \mathbb{P}\{S_a(n) - S_b(n) = s | \mathcal{F}_{S_b}\}\right) \leq \sqrt{\delta} \cdot \max_{b < i \leq a} \left\{ p_e^{\frac{l_i}{p_i}} \right\}$ for $1 \leq b < a \leq \mu$ where $\mathcal{F}_{S_i}$ is the σ-field generated by $\beta[n - S_i(n)], \beta[n - S_i(n) + 1], \cdots, \beta[n]$.

Then, here is the proposed suboptimal maximum likelihood estimator for $\mathbf{x}[n]$:

$$\widehat{\mathbf{x}}[n] = \mathbf{M_n} \begin{bmatrix} \mathbf{y}[n - k_1] \\ \mathbf{y}[n - k_2] \\ \vdots \\ \mathbf{y}[n - k_{\sum_{1 \leq i \leq \mu} m'_i}] \end{bmatrix} \tag{2.41}$$

Here, k_i also depends on n, but we omit the dependency in notation for simplicity. Notice that the number of observations that this estimator uses, $k_{\sum_{1 \leq i \leq \mu} m'_i}$, can be much larger than the dimension of the system, m. In other words, the estimator proposed here may use many more observations than the number of states (the number of observations that a simple matrix inverse observer needs). This is because we use a successive decoding idea in the proof of Lemma 2.3.

• Analysis of the estimation error: Now, we will analyze the performance of the proposed estimator. Recall that p' is defined in (2.40) and $\delta > 1$. By (iv') and (v'), we can find $c > 0$ that satisfies the following four conditions:

(i") $(1 + k^{p'}) \leq c \cdot \delta^k$ for all $k \geq 0$

(ii") $p_i(k) \leq c \cdot \delta^k$ for all $1 \leq i \leq \mu$ and $k \geq 0$

(iii") $\mathbb{P}\{S_i(n) = s\} \leq c \cdot (\delta \cdot \max_{1 \leq j \leq i} \left\{ p_e^{\frac{l_j}{p_j}} \right\})^s$ for all $1 \leq i \leq \mu$ and $s \in \mathbb{Z}^+$

(iv") $\mathbb{P}\{S_a(n) - S_b(n) = s | \mathcal{F}_{S_b}\} \leq c \cdot (\delta \cdot \max_{b < i \leq a} \left\{ p_e^{\frac{l_i}{p_i}} \right\})^s$ for all $1 \leq b < a \leq \mu$ and $s \in \mathbb{Z}^+$.

Let $\mathcal{F}_\beta$ be the σ-field generated by $\beta[n]$. Then, k_i and S_i are deterministic variables conditioned on $\mathcal{F}_\beta$. The estimation error is upper bounded by

$$
\begin{aligned}
\mathbb{E}[|\mathbf{x}[n]-\widehat{\mathbf{x}}[n]|_2^2] &= \mathbb{E}[\mathbb{E}[|\mathbf{x}[n]-\widehat{\mathbf{x}}[n]|_2^2|\mathcal{F}_\beta]] \\
&\overset{(A)}{=} \mathbb{E}[\mathbb{E}[\left|\mathbf{x}[n]-\mathbf{M_n}(\begin{bmatrix}\mathbf{CA}^{-k_1}\\ \mathbf{CA}^{-k_2}\\ \vdots \\ \mathbf{CA}^{-k_{\sum_{1\le i\le\mu} m_i'}}\end{bmatrix}\mathbf{x}[n]-\begin{bmatrix}\mathbf{v}'[n-k_1]\\ \mathbf{v}'[n-k_2]\\ \vdots \\ \mathbf{v}'[n-k_{\sum_{1\le i\le\mu} m_i'}]\end{bmatrix})\right|_2^2|\mathcal{F}_\beta]] \\
&\overset{(B)}{=} \mathbb{E}[\mathbb{E}[\left|\mathbf{M_n}\begin{bmatrix}\mathbf{v}'[n-k_1]\\ \mathbf{v}'[n-k_2]\\ \vdots \\ \mathbf{v}'[n-k_{\sum_{1\le i\le\mu} m_i'}]\end{bmatrix}\right|_2^2|\mathcal{F}_\beta]] \\
&\lesssim \mathbb{E}[|\mathbf{M_n}|_{max}^2\cdot\mathbb{E}[\left|\begin{bmatrix}\mathbf{v}'[n-k_1]\\ \mathbf{v}'[n-k_2]\\ \vdots \\ \mathbf{v}'[n-k_{\sum_{1\le i\le\mu} m_i'}]\end{bmatrix}\right|_{max}^2|\mathcal{F}_\beta]] \\
&\overset{(C)}{\lesssim} \mathbb{E}[|\mathbf{M_n}|_{max}^2\cdot(1+S_\mu^{p'}(n))^2] \\
&\overset{(D)}{\le} \mathbb{E}[\max_{1\le i\le\mu}\{\left(\frac{p_i(S_i(n))}{\epsilon}|\lambda_{i,1}|^{S_i(n)}\right)^2\}\cdot(1+S_\mu^{p'}(n))^2] \\
&\le \sum_{1\le i\le\mu}\mathbb{E}[\left(\frac{p_i(S_i(n))}{\epsilon}|\lambda_{i,1}|^{S_i(n)}\right)^2\cdot(1+S_\mu^{p'}(n))^2] \\
&\overset{(E)}{\lesssim} \sum_{1\le i\le\mu}\mathbb{E}[\delta^{2S_i(n)}\cdot|\lambda_{i,1}|^{2S_i(n)}\cdot\delta^{2S_\mu(n)}] \\
&= \sum_{1\le i\le\mu}\mathbb{E}[\delta^{4S_i(n)}\cdot|\lambda_{i,1}|^{2S_i(n)}\cdot\mathbb{E}[\delta^{2(S_\mu(n)-S_i(n))}|\mathcal{F}_{S_i(n)}]] \\
&\overset{(F)}{\lesssim} \sum_{1\le i\le\mu}\mathbb{E}[\delta^{4S_i(n)}\cdot|\lambda_{i,1}|^{2S_i(n)}\cdot\sum_{s=0}^{\infty}\delta^{2s}\cdot(\delta\cdot\max_{1\le i\le\mu}\{p_e^{\frac{l_i}{p_i}}\})^s] \\
&\overset{(G)}{\lesssim} \sum_{1\le i\le\mu}\mathbb{E}[\delta^{4S_i(n)}\cdot|\lambda_{i,1}|^{2S_i(n)}] \\
&\overset{(H)}{\lesssim} \sum_{1\le i\le\mu}\sum_{s=0}^{\infty}\delta^{4s}\cdot|\lambda_{i,1}|^{2s}\cdot(\delta\cdot\max_{1\le j\le i}\left\{p_e^{\frac{l_j}{p_j}}\right\})^s \\
&= \sum_{1\le i\le\mu}\sum_{s=0}^{\infty}(\delta^5\cdot|\lambda_{i,1}|^2\cdot\max_{1\le j\le i}\left\{p_e^{\frac{l_j}{p_j}}\right\})^s \\
&\overset{(I)}{<} \infty
\end{aligned}
$$

where $\lesssim$ holds for all n.

(A): By (2.39) and (2.41).

(B): By condition (ii').

(C): Since $\mathbb{E}[\mathbf{v}'[n-k]^\dagger \mathbf{v}'[n-k]] \lesssim 1+k^{p'}$ by the definition of p' of (2.40), and thus each element of the $\mathbf{v}'[n]$ vector obeys max bound.

(D): By condition (iii').

(E): By condition (i") and (ii").

(F): By condition (iv").

(G): Since $\delta^5 \cdot \max_{1 \leq i \leq \mu} p_e^{\frac{l_i}{p_i}} |\lambda_{i,1}|^2 < 1$.

(H): By condition (iii").

(I): Since $\delta^5 \cdot \max_{1 \leq i \leq \mu} p_e^{\frac{l_i}{p_i}} |\lambda_{i,1}|^2 < 1$.

Therefore, the estimation error variance is uniformly bounded over n when $p_e < \frac{1}{\max_{1 \leq i \leq \mu} |\lambda_{i,1}|^{2\frac{p_i}{l_i}}}$, which finishes the proof.

2.8.4 Necessity Proof of Theorem 2.7

Intuitively, we will give all states except the ones that corresponds to the bottleneck eigenvalue cycle as side-information to the estimator. Then, the problem reduces to the single eigenvalue cycle one discussed in Section 2.5.1, and we can prove the estimation error diverges similarly. This argument works for $p_e > \frac{1}{\max_i |\lambda_{i,1}|^{2\frac{p_i}{l_i}}}$, since we can show that a single additional disturbance $\mathbf{w}[n]$ grows exponentially. However, for the equality case $p_e = \frac{1}{\max_i |\lambda_{i,1}|^{2\frac{p_i}{l_i}}}$, the proof can be more complicated since not a single disturbance but the sum of disturbances linearly diverges to infinity.

So, to make this argument complete and rigorous, we will analyze the optimal estimator, and prove that its estimation error diverges when the condition of the lemma is violated.

It is well-known that the optimal estimator is the Kalman filter and it can be written in recursive form. Let $\mathcal{F}_\beta$ be the σ-field generated by $\beta[n]$. Denote the one-step prediction error as $\mathbf{\Sigma_{n+1|n}} := \mathbb{E}[(\mathbf{x}[n+1] - \mathbb{E}[\mathbf{x}[n+1]|\mathbf{y}^n])(\mathbf{x}[n+1] - \mathbb{E}[\mathbf{x}[n+1]|\mathbf{y}^n])^\dagger | \mathcal{F}_\beta]$. Then, $\mathbf{\Sigma_{n+1|n}}$ follows the following recursive equation [55, p.101].

$$\mathbf{\Sigma_{n+1|n}} = (\mathbf{A} - \mathbf{A}\mathbf{L_n}\mathbf{\bar{C}_n})\mathbf{\Sigma_{n|n-1}}(\mathbf{A} - \mathbf{A}\mathbf{L_n}\mathbf{\bar{C}_n})^\dagger + \mathbf{A}\mathbf{L_n}\mathbb{E}[\mathbf{v}[n]\mathbf{v}[n]^\dagger]\mathbf{L_n}^\dagger\mathbf{A}^\dagger + \mathbf{B}\mathbb{E}[\mathbf{w}[n]\mathbf{w}[n]^\dagger]\mathbf{B}^\dagger \tag{2.42}$$

Here, $\mathbf{L_n} = \mathbf{\Sigma_{n|n-1}}\mathbf{\bar{C}_n^\dagger}\left[\mathbf{\bar{C}_n}\mathbf{\Sigma_{n|n-1}}\mathbf{\bar{C}_n^\dagger} + \mathbb{E}[\mathbf{v}[n]\mathbf{v}[n]^\dagger]\right]^{-1}$, and $\mathbf{\bar{C}_n} = \mathbf{C}$ if $\beta[n] = 1$ and $\mathbf{\bar{C}_n} = \mathbf{0}$ otherwise. Notice that $\mathbf{\Sigma_{n+1|n}}$ is a random variable.

Moreover, it is also known that when $(\mathbf{A}, \mathbf{B})$ is controllable, the one-step prediction error of $\mathbf{x}[n+1]$ based on $\mathbf{y}[n]$ becomes positive definite for large enough n even if there are no erasures. Therefore, there exists $m \in \mathbb{N}$ and $\sigma^2 > 0$ such that $\mathbf{\Sigma_{n+1|n}} \succeq \sigma^2\mathbf{I}$ with probability one for all

$n \geq m$. Therefore, by (2.42) for all $n \geq n' \geq m$ we have

$$\begin{aligned}\mathbf{\Sigma_{n+1|n}} &\succeq (\mathbf{A} - \mathbf{A}\mathbf{L_n}\bar{\mathbf{C}}_\mathbf{n}) \cdots (\mathbf{A} - \mathbf{A}\mathbf{L_{n'}}\bar{\mathbf{C}}_\mathbf{n'})\mathbf{\Sigma_{n'|n'-1}}(\mathbf{A} - \mathbf{A}\mathbf{L_{n'}}\bar{\mathbf{C}}_\mathbf{n'})^\dagger \cdots (\mathbf{A} - \mathbf{A}\mathbf{L_n}\bar{\mathbf{C}}_\mathbf{n})^\dagger \\ &\succeq \sigma^2(\mathbf{A} - \mathbf{A}\mathbf{L_n}\bar{\mathbf{C}}_\mathbf{n}) \cdots (\mathbf{A} - \mathbf{A}\mathbf{L_{n'}}\bar{\mathbf{C}}_\mathbf{n'})\mathbf{I}(\mathbf{A} - \mathbf{A}\mathbf{L_{n'}}\bar{\mathbf{C}}_\mathbf{n'})^\dagger \cdots (\mathbf{A} - \mathbf{A}\mathbf{L_n}\bar{\mathbf{C}}_\mathbf{n})^\dagger. \end{aligned} \quad (2.43)$$

Let's use the definitions of $\mathbf{U}$, $\mathbf{A'}$, $\mathbf{C'}$, $\mathbf{U}$, $\mathbf{A_i}$, $\mathbf{C_i}$, $\lambda_{i,j}$, p_i, l_i, ν_i from (2.16), (2.17) and (2.18). Let $i^\star := \underset{1 \leq i \leq \mu}{argmax} |\lambda_{i,1}|^{2\frac{p_i}{l_i}}$. Let $S'^\star \subseteq \{0, 1, \cdots, p_{i^\star} - 1\}$ be a set achieving the minimum cardinality $l_{i^\star}$. In other words, define $S^\star := \{s_1^\star, s_2^\star, \cdots, s_{|S^\star|}^\star\} = \{0, 1, \cdots, p_{i^\star} - 1\} \setminus S'^\star$. Then, $|S'^\star| = l_{i^\star}$ and

$$\begin{bmatrix} \mathbf{C_{i^\star}} \mathbf{A_{i^\star}}^{s_1^\star} \\ \mathbf{C_{i^\star}} \mathbf{A_{i^\star}}^{s_2^\star} \\ \vdots \\ \mathbf{C_{i^\star}} \mathbf{A_{i^\star}}^{s_{|S^\star|}^\star} \end{bmatrix}$$

is rank deficient, i.e. the rank is strictly less than $\nu_{i^\star}$.

For a given time index n, define the stopping time S_n as the most recent observation which does not belong to $S^\star$ in modulo $p_{i^\star}$, i.e.

$$\begin{aligned} S_n := \inf\{kp_{i^\star} : k \in \mathbb{Z}^+ &\text{ and there exists } k' \text{ such that} \\ &\beta[n-k'] = 1, kp_{i^\star} \leq k' < (k+1)p_{i^\star}, -k' - 1 (mod\ p_{i^\star}) \in S'^\star\}. \end{aligned}$$

Then, we can compute that $\mathbb{P}\{S_n = kp_{i^\star}\} = (1 - p_e^{l_{i^\star}})(p_e^{l_{i^\star}})^k$ for all $k \in \mathbb{Z}^+$. From the definition of S_n, we can see that for all $0 \leq k < S_n$, $\beta[n-k] = 1$ only if $-k - 1(mod\ p_{i^\star}) \in S^\star$.

Then, conditioned on $n - S_n \geq m$, by (2.43) the following inequality holds with probability one:

$$\mathbf{\Sigma_{n+1|n}} \succeq \sigma^2(\mathbf{A} - \mathbf{A}\mathbf{L_n}\bar{\mathbf{C}}_\mathbf{n}) \cdots (\mathbf{A} - \mathbf{A}\mathbf{L_{n-S_n+1}}\bar{\mathbf{C}}_\mathbf{n-S_n+1})\mathbf{I}(\mathbf{A} - \mathbf{A}\mathbf{L_{n-S_n+1}}\bar{\mathbf{C}}_\mathbf{n-S_n+1})^\dagger \cdots (\mathbf{A} - \mathbf{A}\mathbf{L_n}\bar{\mathbf{C}}_\mathbf{n})^\dagger. \quad (2.44)$$

where $\bar{\mathbf{C}}_\mathbf{n-S_n+k} = \mathbf{C}$ or $\mathbf{0}$ if $-S_n + k - 1(mod\ p_{i^\star}) = k - 1(mod\ p_{i^\star}) \in S^\star$ and $\bar{\mathbf{C}}_\mathbf{n-S_n+k} = \mathbf{0}$ if $-S_n + k - 1(mod\ p_{i^\star}) = k - 1(mod\ p_{i^\star}) \in S'^\star$. Here, $-S_n + k - 1(mod\ p_{i^\star}) = k - 1(mod\ p_{i^\star})$ follows from that $S_n(mod\ p_{i^\star}) = 0$ by the definition of S_n.

We will prove that the L.H.S. of (2.44) grows exponentially. For this, we first need the following lemma.

Lemma 2.4. *Consider* $\mathbf{A}$, $\mathbf{C}$, $\mathbf{U}$, $\mathbf{A'}$, $\mathbf{C'}$, $\mathbf{A_i}$, $\mathbf{C_i}$, ν_i, p_i *given in* (2.16), (2.17) *and* (2.18). *For a given set* $S := \{s_1, \cdots, s_{|S|}\} \subseteq \{0, 1, \cdots, p_i - 1\}$, *let* $\begin{bmatrix} \mathbf{C_i}\mathbf{A_i}^{s_1} \\ \mathbf{C_i}\mathbf{A_i}^{s_2} \\ \vdots \\ \mathbf{C_i}\mathbf{A_i}^{s_{|S|}} \end{bmatrix}$ *be rank-deficient, i.e. the rank is* *less than ν_i, and define*

$$\bar{\mathbf{A}}(\mathbf{K_0}, \cdots, \mathbf{K_{p_i-1}}) := (\mathbf{A} - \mathbf{K_{p_i-1}}\bar{\mathbf{C}}_{\mathbf{p_i-1}}) \cdots (\mathbf{A} - \mathbf{K_0}\bar{\mathbf{C}}_\mathbf{0})$$

where $\bar{\mathbf{C}}_\mathbf{j} = \mathbf{C}$ or $\mathbf{0_{l\times m}}$ when $j \in S$ and $\bar{\mathbf{C}}_\mathbf{j} = \mathbf{0_{l\times m}}$ otherwise.
Then, for all $\mathbf{K_0}, \cdots, \mathbf{K_{p_i-1}} \in \mathbb{C}^{m\times l}$, $\bar{\mathbf{A}}(\mathbf{K_0}, \cdots, \mathbf{K_{p_i-1}})$ has a common right eigenvector $\mathbf{e}$ whose eigenvalue is $\lambda_{i,1}^{p_i}$.

Proof. For simplicity of notation, we will set $i = 1$, but the proof for general i is the same. Let $\mathbf{e}' = \begin{bmatrix} e_1 \\ \vdots \\ e_{\nu_1} \end{bmatrix}$ be a nonzero vector that belongs to the right null space of $\begin{bmatrix} \mathbf{C_1}\mathbf{A_1}^{s_1} \\ \mathbf{C_1}\mathbf{A_1}^{s_2} \\ \vdots \\ \mathbf{C_1}\mathbf{A_1}^{s_{|S|}} \end{bmatrix}$. Let $\mathbf{e_1'}$ be a $m_{1,1} \times 1$ column vector whose first element is e_1 and the rest are 0. Likewise, $\mathbf{e_2'}$ is a $m_{1,2} \times 1$ column vector with first element e_2 and the rest 0. $\mathbf{e_3'}, \cdots, \mathbf{e'_{\nu_1}}$ are defined in the same way. Let a $m \times 1$ column vector $\mathbf{e}''$ be $\begin{bmatrix} \mathbf{e_1'} \\ \vdots \\ \mathbf{e'_{\nu_1}} \\ \mathbf{0}_{(m - \sum_{1\le i\le \nu_1} m_{1,i})\times 1} \end{bmatrix}$. Then, we will prove that $\mathbf{e} := \mathbf{U}\mathbf{e}''$ is the eigenvector that satisfies the conditions of the lemma.

By construction, we can see that $\mathbf{C_1}\mathbf{A_1}^k\mathbf{e}' = \mathbf{0}$ for $k \in \{s_1, \cdots, s_{|S|}\}$. Moreover, since $\mathbf{C}\mathbf{A^k}\mathbf{e} = \mathbf{C}\mathbf{U}\mathbf{A}'^k\mathbf{U}^{-1}\mathbf{U}\mathbf{e}'' = \mathbf{C}'\mathbf{A}'^k\mathbf{e}''$, we also have $\mathbf{C}\mathbf{A}^k\mathbf{e} = \mathbf{0}$ for $k \in \{s_1, \cdots, s_{|S|}\}$. Thus, we can conclude

$$\begin{aligned}
&(\mathbf{A} - \mathbf{K_{p_1-1}}\bar{\mathbf{C}}_{\mathbf{p_1-1}}) \cdots (\mathbf{A} - \mathbf{K_{s_1}}\bar{\mathbf{C}}_{\mathbf{s_1}})(\mathbf{A} - \mathbf{K_{s_1-1}}\bar{\mathbf{C}}_{\mathbf{s_1-1}}) \cdots (\mathbf{A} - \mathbf{K_0}\bar{\mathbf{C}}_\mathbf{0})\mathbf{e} \\
&= (\mathbf{A} - \mathbf{K_{p_1-1}}\bar{\mathbf{C}}_{\mathbf{p_1-1}}) \cdots (\mathbf{A} - \mathbf{K_{s_1}}\bar{\mathbf{C}}_{\mathbf{s_1}})(\mathbf{A} - \mathbf{K_{s_1-1}}\mathbf{0}) \cdots (\mathbf{A} - \mathbf{K_0}\mathbf{0})\mathbf{e} \\
&= (\mathbf{A} - \mathbf{K_{p_1-1}}\bar{\mathbf{C}}_{\mathbf{p_1-1}}) \cdots (\mathbf{A} - \mathbf{K_{s_1}}\bar{\mathbf{C}}_{\mathbf{s_1}})\mathbf{A}^{s_1}\mathbf{e} \\
&= (\mathbf{A} - \mathbf{K_{p_1-1}}\bar{\mathbf{C}}_{\mathbf{p_1-1}}) \cdots (\mathbf{A}^{s_1+1}\mathbf{e} - \mathbf{K_{s_1}}\bar{\mathbf{C}}_{\mathbf{s_1}}\mathbf{A}^{s_1}\mathbf{e}) \\
&\overset{(a)}{=} (\mathbf{A} - \mathbf{K_{p_1-1}}\bar{\mathbf{C}}_{\mathbf{p_1-1}}) \cdots (\mathbf{A}^{s_1+1}\mathbf{e}) \\
&\overset{(b)}{=} \mathbf{A}^{p_1}\mathbf{e} = \mathbf{U}\mathbf{A}'^{p_1}\mathbf{U}^{-1}\mathbf{e} = \mathbf{U}\mathbf{A}'^{p_1}\mathbf{e}'' \\
&\overset{(c)}{=} \mathbf{U}\lambda_{1,1}^{p_1}\mathbf{e}'' = \lambda_{1,1}^{p_1}\mathbf{e}
\end{aligned}$$

(a): $\mathbf{C}\mathbf{A}^{s_1}\mathbf{e} = \mathbf{0}$ and $\mathbf{0} \cdot \mathbf{A}^{s_1}\mathbf{e} = \mathbf{0}$.
(b): Repetitive use of (a) for $s_2, \cdots, s_{|S|}$.
(c): $\mathbf{A_1}^{p_1} = \lambda_{1,1}^{p_1}\mathbf{I}$ and the definition of the vector $\mathbf{e}''$.

Thus, the lemma is proved. □

Let the vector $\mathbf{e}$ be the right eigenvector of Lemma 2.4 for $i = i^\star$. Then, there exists $\sigma' > 0$

such that $\mathbf{I} \succeq \sigma'^2 \mathbf{e}\mathbf{e}^\dagger$. (2.44) is lower bounded as

$$\mathbf{\Sigma_{n+1|n}} \succeq \sigma^2 \sigma'^2 \lambda_{i^\star,1}^{S_n} \mathbf{e}\mathbf{e}^\dagger (\lambda_{i^\star,1}^{S_n})^\dagger.$$

Since $p_e \geq \frac{1}{|\lambda_{i^\star,1}|^{2\frac{p_{i^\star}}{l_{i^\star}}}}$, the expected one-step prediction error is lower bounded as follows:[16]

$$\begin{aligned}
&\mathbb{E}[(\mathbf{x}[n+1] - \mathbb{E}[\mathbf{x}[n+1]|\mathbf{y}^n])^\dagger(\mathbf{x}[n+1] - \mathbb{E}[\mathbf{x}[n+1]|\mathbf{y}^n])] \\
&\geq \mathbb{E}[\sigma^2\sigma'^2 |\lambda_{i^\star,1}|^{2S_n} |\mathbf{e}|^2 \cdot \mathbf{1}(n - S_n \geq m)] \\
&\geq \sigma^2\sigma'^2|\mathbf{e}|^2 \sum_{0 \leq s \leq \lfloor \frac{n-m}{p_{i^\star}} \rfloor} (1 - p_e^{l_{i^\star}})(|\lambda_{i^\star,1}|^{2p_{i^\star}} p_e^{l_{i^\star}})^s \\
&\geq \sigma^2\sigma'^2|\mathbf{e}|^2 \cdot (1 - p_e^{l_{i^\star}}) \cdot (\lfloor \frac{n-m}{p_{i^\star}} \rfloor).
\end{aligned}$$

Therefore, as n goes to infinity, the one-step prediction error diverges to infinity. The estimation error variance for the state is not uniformly bounded either, so the system is not intermittent observable.

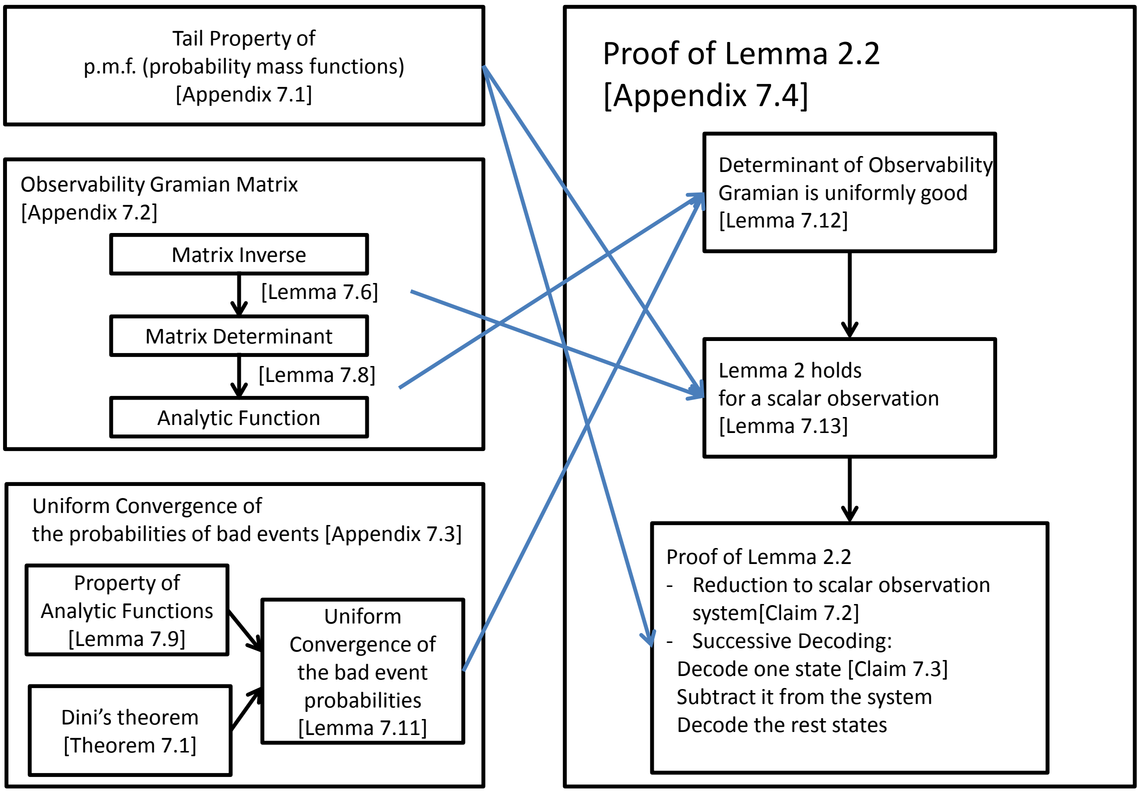


Figure 2.5: Flow diagram of the proof of Lemma 2.2

[16]The lower bound does not hold when $|\lambda_{i^\star,1}| = 1$ which induces $p_e = 1$. However, in this case we do not have any observations, so trivially the system is unstable.

2.8.5 Proof Outline of Lemma 2.2 and Lemma 2.3

Now, the proofs of Theorem 2.8 and 2.7 boil down to the proofs of Lemma 2.2 and 2.3. Since the proofs of Lemma 2.2 and 2.3 shown in Appendix are too involved, we give the outlines of the proofs in this section.

Proof Outline of Lemma 2.2

The proof flow of Lemma 2.2 is shown in Figure 2.5. As we discussed in Section 2.5 by explicit examples and as we formally saw in Section 2.8.1, the sufficiency proof of the critical erasure probability mainly relies on two mathematical notions, the p.m.f. (probability mass function) tail of stopping times and the observability gramian of the system.

In Appendix 7.1, we first study some well-known properties of the p.m.f. tail of random variables which will be used to model the stopping times of interest. For example, we will prove that when we add two independent random variables, the p.m.f. tail of the resulting random variable is decided by the thicker one.

In Appendix 7.2, we consider the second notion: the observability gramian of the system. We used a sub-optimal maximum likelihood estimator in the sufficiency proof of Section 2.8.1, and its performance heavily relies on this inverse matrix of the observability gramian, especially the norm of the inverse matrix. However, the norm of a matrix depends on all elements of the matrix, and so it is hard to compute. Instead, we first relate the norm of the matrix with the determinant of the matrix in Lemma 7.6. Thus, we can focus on the determinant of the observability gramian instead of the norm to analyze the performance of the estimator. Furthermore, the determinant of the observability gramian is an analytic function of the sampling times. Therefore, in Lemma 7.8 we will further reduce the estimator performance problem to a question about analytic functions. More precisely, we will prove that when the relevant analytic functions are large enough, the proposed maximum likelihood estimator performs well.

Now, we can focus on a particular set of analytic functions. We have to prove that after introducing nonuniform sampling, the multiple analytic functions which reflect different erasure patterns (observation time indexes) are uniformly large enough with high probability. In Appendix 7.3, we will prove that the probabilities that the relevant analytic functions are too small converge to zero uniformly over all erasure patterns.

To show this, we first start with a single analytic function. In Lemma 7.9, we will prove that for a given erasure pattern, each relevant analytic function is large enough with high probability after nonuniform sampling. To convert this pointwise convergence result to a uniform convergence result, we will use Dini's theorem [35]. Dini's theorem assures that under compactness and monotone convergence conditions, pointwise convergence implies uniform convergence. Using these facts, Lemma 7.11 proves the desired uniform convergence, i.e. the relevant set of analytic functions are uniformly large enough with high probability for all erasure patterns.

Now, we are ready to prove the desired Lemma 2.2. In Appendix 7.2, we relate the norm of the inverse matrix of the observability gramian to an analytic function. In Appendix 7.3, we found that these analytic functions uniformly converge. Thus, by integrating these results, Lemma 7.12 shows that the norm of the inverse matrix of the observability gramian is large enough with high probability uniformly over all erasure patterns. In Section 2.5, we saw that erasure patterns are modeled by geometric random variables. Now, we can apply the p.m.f. tail properties studied in Appendix 7.1 to understand the p.m.f. tail of these geometric random variables. With this understanding, we can easily prove Lemma 7.13, which tells us that Lemma 2.2 holds for systems with scalar observations.

Finally, the only remaining step is generalizing this fact to systems with vector observations. For this, we adapt an induction argument and use successive decoding ideas [21]. Induction is on the number of states of the plant. First, in Claim 7.2 we reduce the system with vector observations to another system with scalar observations by multiplying a proper row vector to the observations. Then, as shown in Claim 7.3, we can apply the result for systems with scalar observations to estimate just one particular state. Once we estimate one state, we can remove the estimated state from the system to get a new system with a smaller number of states. Here, this idea of estimating a part of state and subtracting it from the original system is known as successive decoding in the information theory community [21]. Since we now have a system with a smaller number of states, we can apply the induction hypothesis to finish the proof of Lemma 2.2.

Let's consider the following system to understand this last step more precisely.

$$\begin{bmatrix} x_1[n+1] \\ x_2[n+1] \\ x_3[n+1] \\ x_4[n+1] \\ x_5[n+1] \end{bmatrix} = \begin{bmatrix} 3 & 1 & & & \\ & 3 & & & \\ & & 3 & 1 & \\ & & & 3 & \\ & & & & 2 \end{bmatrix} \begin{bmatrix} x_1[n] \\ x_2[n] \\ x_3[n] \\ x_4[n] \\ x_5[n] \end{bmatrix} + \begin{bmatrix} w_1[n] \\ w_2[n] \\ w_3[n] \\ w_4[n] \\ w_5[n] \end{bmatrix}$$

$$\mathbf{y}[n] = \beta[n] \left(\begin{bmatrix} 1 & 3 & 1 & 2 & 1 \\ 1 & 2 & -1 & 3 & 1 \end{bmatrix} \mathbf{x}[n] \right)$$

As we can see, the above system has vector observations. To reduce it to a scalar observation system, we multiply a row vector $\begin{bmatrix} 1 & -1 \end{bmatrix}$ to each observation. Then, the resulting scalar observation becomes

$$\begin{bmatrix} 1 & -1 \end{bmatrix} \mathbf{y}[n] = \beta[n] \begin{bmatrix} 0 & 1 & 2 & -1 & 0 \end{bmatrix} \mathbf{x}[n].$$

Here, with new scalar observations, the states $x_1[n]$ and $x_5[n]$ are unobservable (unobservability of $x_1[n]$ is the key in the following argument). Thus, we will reduce the system to the following observable system by considering a function of states as a new state.

$$\begin{bmatrix} \frac{1}{2}x_2[n+1] + x_3[n+1] \\ x_4[n+1] \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ & 3 \end{bmatrix} \begin{bmatrix} \frac{1}{2}x_2[n] + x_3[n] \\ x_4[n] \end{bmatrix} + \begin{bmatrix} \frac{1}{2}w_2[n] + w_3[n] \\ w_4[n] \end{bmatrix}$$

$$\begin{bmatrix} 1 & -1 \end{bmatrix} \mathbf{y}[n] = \beta[n] \begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2}w_2[n] + w_3[n] \\ w_4[n] \end{bmatrix}$$

By considering $\frac{1}{2}x_2[n] + x_3[n]$ as one state, the resulting system becomes an observable scalar-observation system and the state $x_4[n]$ remains intact. This step is what Claim 7.2 does.

Therefore, using the result about observable scalar-observation systems, we can first estimate $x_4[n]$ (this step corresponds to Claim 7.3). Once we have an estimate $\hat{x}_4[n]$, we can subtract it from the observation.

$$\begin{aligned} \mathbf{y}[n] - \beta[n] \begin{bmatrix} 2 \\ 3 \end{bmatrix} \hat{x}_4[n] &= \beta[n] \left(\begin{bmatrix} 1 & 3 & 1 & 2 & 1 \\ 1 & 2 & -1 & 3 & 1 \end{bmatrix} \mathbf{x}[n] - \begin{bmatrix} 2 \\ 3 \end{bmatrix} x_4[n] \right) \\ &= \beta[n] \left(\begin{bmatrix} 1 & 3 & 1 & 1 \\ 1 & 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1[n] \\ x_2[n] \\ x_3[n] \\ x_5[n] \end{bmatrix} + \begin{bmatrix} 2 \\ 3 \end{bmatrix} x_4[n] - \begin{bmatrix} 2 \\ 3 \end{bmatrix} \hat{x}_4[n] \right) \end{aligned}$$

Therefore, the resulting system can be thought as of a system with only four states (one state less than the original system). Using induction hypothesis, we can estimate the remaining four states.

Proof Outline of Lemma 2.3

The main proof ideas and structure of Lemma 2.3 described in Figure 2.6 are essentially the same as those of Lemma 2.2. Thus, here we will mainly emphasize the differences between the two proofs.

We reuse the properties of the p.m.f. tails that we proved in Appendix 7.1. In Appendix 7.5, we relate the norm of the observability gramian inverse matrix with an analytic functions just as we did in Appendix 7.2.

In Appendix 7.6, we will essentially show that the relevant set of analytic functions are uniformly large enough for almost all erasure patterns. However, there is a crucial difference from the nonuniform sampling case of Appendix 7.3. Unlike the nonuniform sampling case, there is no randomness which jitters the sampling time. Therefore, we have to count the number of erasure patterns which make the relevant analytic functions small (rather than computing a probability), and prove that the number of such patterns is small enough compared to the number of all possible erasure patterns.

For this, we use Weyl's criterion [54] which gives us a handle on the ergodic behavior of sequences. Specifically, a sequence $\alpha - \lfloor \alpha \rfloor, 2\alpha - \lfloor 2\alpha \rfloor, 3\alpha - \lfloor 3\alpha \rfloor, \cdots$ with irrational α can be modeled by a uniform random variable on $[0, 1]$. Therefore, using this fact we can reduce the counting problem

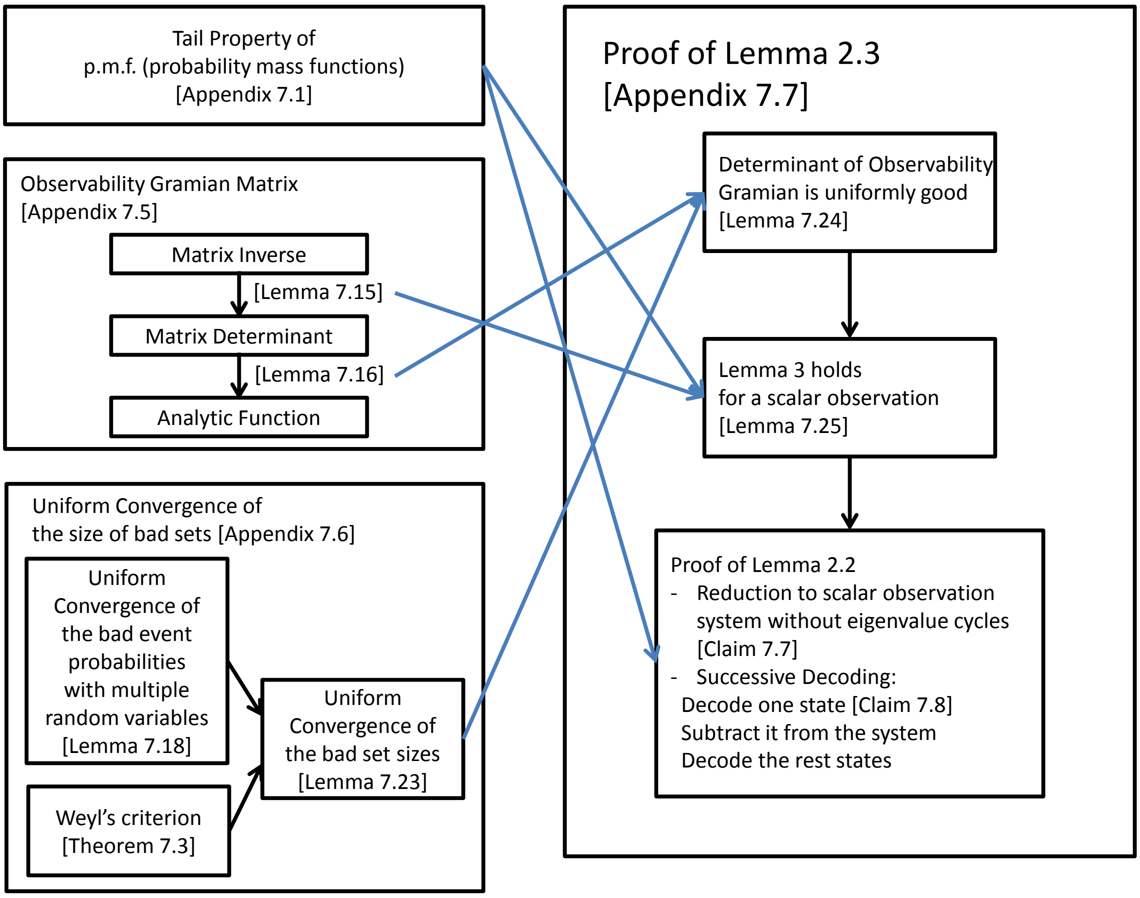


Figure 2.6: Flow diagram of the proof of Lemma 2.3

that we are facing to basically the same probability problem in the spirit of one we already studied in Appendix 7.3. However, there is still a difference between these two cases. We may need multiple random variables to model the erasure sequences. To clarify this point, let's consider the following examples.

Let $\mathbf{A_1} = \begin{bmatrix} e^{j\sqrt{2}} & 0 \\ 0 & e^{j2\sqrt{2}} \end{bmatrix}$, $\mathbf{A_2} = \begin{bmatrix} e^{j\sqrt{2}} & 0 \\ 0 & e^{j\sqrt{3}} \end{bmatrix}$, $\mathbf{C} = \begin{bmatrix} 1 & 1 \end{bmatrix}$. The row of the observability gramian of $(\mathbf{A_1}, \mathbf{C})$ is $\mathbf{C}\mathbf{A_1}^n = \begin{bmatrix} e^{j\sqrt{2}n} & e^{j2\sqrt{2}n} \end{bmatrix}$. In this case, the elements of $\mathbf{C}\mathbf{A_1}^n$ do not satisfy Weyl's criterion [54]. It can be approximated by $\begin{bmatrix} e^{jX} & e^{j2X} \end{bmatrix}$ where X is uniform in $[0, 2\pi]$ — so it involves only one random variable.

However, the row of the observability gramian of $(\mathbf{A_2}, \mathbf{C})$ is $\mathbf{C}\mathbf{A_2}^n = \begin{bmatrix} e^{j\sqrt{2}n} & e^{j\sqrt{3}n} \end{bmatrix}$ whose elements satisfy Weyl's criterion [54]. Thus, it can be approximated by $\begin{bmatrix} e^{jX_1} & e^{jX_2} \end{bmatrix}$ where X_1, X_2 are independent uniform random variables in $[0, 2\pi]$ — so it involves two random variables.

Therefore, in Lemma 7.18, we first extend the results of Appendix 7.3 to multiple random variables. Then, by combining Lemma 7.18 with Weyl's criterion, Lemma 7.23 shows the number of bad erasure patterns that make the relevant analytic functions small is small enough uniformly over all the analytic functions.

In Appendix 7.7, we are finally prove Lemma 2.3. First, we prove the lemma for systems with scalar observations and without eigenvalue cycles. Lemma 7.24 and Lemma 7.25 parallel Lemma 7.12 and Lemma 7.13. Thus, the final step is extending the result to general systems with vector observations and with eigenvalue cycles.

The proof ideas are similar to those of Lemma 2.2 except that there is another difficulty of handling eigenvalue cycles. The main ideas for the proof are still induction and successive decoding. However, we also adapt polyphase decomposition ideas from digital signal processing [75] to handle the eigenvalue cycles. More precisely, by sub-sampling systems by the period of the system, we decompose one periodic system (with eigenvalue cycles) to multiple aperiodic systems (without eigenvalue cycles). Using these ideas, we can reduce the original system with vector observations and eigenvalue cycles to multiple sub-sampled systems with scalar observations and without eigenvalue cycles. Then, we can decode one state out of the reduced sub-sampled systems. We can apply successive decoding ideas and the induction to finish the proof of Lemma 2.3.

These ideas can be clarified by the following descriptive example. Consider the following system with eigenvalue cycles and vector observations.

$$\begin{bmatrix} x_1[n+1] \\ x_2[n+1] \\ x_3[n+1] \\ x_4[n+1] \end{bmatrix} = \begin{bmatrix} 3 & & & \\ & 3 & & \\ & & -3 & \\ & & & 2 \end{bmatrix} \begin{bmatrix} x_1[n] \\ x_2[n] \\ x_3[n] \\ x_4[n] \end{bmatrix} + \begin{bmatrix} w_1[n] \\ w_2[n] \\ w_3[n] \\ w_4[n] \end{bmatrix}$$

$$\mathbf{y}[n] = \beta[n] \left(\begin{bmatrix} 1 & 1 & 1 & 2 \\ 1 & -1 & 2 & 3 \end{bmatrix} \right)$$

Here, following the notations of (2.35) and (2.36), we can see that $\mathbf{C_1} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \end{bmatrix}$ and $\mathbf{A_1} = \begin{bmatrix} 3 & & \\ & 3 & \\ & & -3 \end{bmatrix}$. Thus, we have

$$\begin{bmatrix} -\frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{C_1} \\ \mathbf{C_1 A_1} \end{bmatrix} = \begin{bmatrix} -\frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \\ 3 & 3 & -3 \\ 3 & -3 & -6 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}.$$

Now, we want to reduce the original system with vector observations and eigenvalue cycles to the one with scalar observations and no eigenvalue cycles. For this, we multiply $\begin{bmatrix} -\frac{1}{4} & \frac{1}{2} \end{bmatrix}$ to even time observations and $\begin{bmatrix} \frac{1}{4} & 0 \end{bmatrix}$ to odd time observations.

$$\begin{bmatrix} x_1[2(n+1)] \\ x_2[2(n+1)] \\ x_3[2(n+1)] \\ x_4[2(n+1)] \end{bmatrix} = \begin{bmatrix} 9 & & & \\ & 9 & & \\ & & 9 & \\ & & & 4 \end{bmatrix} \begin{bmatrix} x_1[2n] \\ x_2[2n] \\ x_3[2n] \\ x_4[2n] \end{bmatrix} + \begin{bmatrix} 3w_1[2n] + w_1[2n+1] \\ 3w_2[2n] + w_2[2n+1] \\ -3w_3[2n] + w_3[2n+1] \\ 2w_4[2n] + w_4[2n+1] \end{bmatrix}$$

$$\begin{bmatrix} -\frac{1}{4} & \frac{1}{2} \end{bmatrix} \mathbf{y}[2n] = \beta[2n] \left(\begin{bmatrix} \frac{1}{4} & -\frac{3}{4} & \frac{3}{4} & 1 \end{bmatrix} \begin{bmatrix} x_1[2n] \\ x_2[2n] \\ x_3[2n] \\ x_4[2n] \end{bmatrix} \right)$$

$$\begin{bmatrix} \frac{1}{4} & 0 \end{bmatrix} \mathbf{y}[2n+1] = \beta[2n+1] \left(\begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} x_1[2n+1] \\ x_2[2n+1] \\ x_3[2n+1] \\ x_4[2n+1] \end{bmatrix} \right)$$

$$= \beta[2n+1] \left(\begin{bmatrix} \frac{3}{4} & \frac{3}{4} & \frac{-3}{4} & 1 \end{bmatrix} \begin{bmatrix} x_1[2n] \\ x_2[2n] \\ x_3[2n] \\ x_4[2n] \end{bmatrix} + \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} w_1[2n] \\ w_2[2n] \\ w_3[2n] \\ w_4[2n] \end{bmatrix} \right)$$

We can consider even-time observations and odd-time observations as two separate systems. Then, these two sub-sampled systems with scalar observations can be rewritten as follows: The first system is

$$\begin{bmatrix} \frac{1}{4}x_1[2(n+1)] - \frac{3}{4}x_2[2(n+1)] + \frac{3}{4}x_3[2(n+1)] \\ x_4[2(n+1)] \end{bmatrix} = \begin{bmatrix} 9 & \\ & 4 \end{bmatrix} \begin{bmatrix} \frac{1}{4}x_1[2n] - \frac{3}{4}x_2[2n] + \frac{3}{4}x_3[2n] \\ x_4[2n] \end{bmatrix}$$

$$+ \begin{bmatrix} \frac{1}{4}(3w_1[2n] + w_1[2n+1]) - \frac{3}{4}(3w_2[2n] + w_2[2n+1]) + \frac{3}{4}(-3w_3[2n] + w_3[2n+1]) \\ 2w_4[2n] + w_4[2n+1] \end{bmatrix}$$

$$\begin{bmatrix} -\frac{1}{4} & \frac{1}{2} \end{bmatrix} \mathbf{y}[2n] = \beta[2n] \left(\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{4}x_1[2n] - \frac{3}{4}x_2[2n] + \frac{3}{4}x_3[2n] \\ x_4[2n] \end{bmatrix} \right)$$

where $\frac{1}{4}x_1[2n] - \frac{3}{4}x_2[2n] + \frac{3}{4}x_3[2n]$ and $x_4[2n]$ are the states of the system.

The second system is

$$\begin{bmatrix} \frac{3}{4}x_1[2(n+1)] + \frac{3}{4}x_2[2(n+1)] - \frac{3}{4}x_3[2(n+1)] \\ x_4[2(n+1)] \end{bmatrix} = \begin{bmatrix} 9 & \\ & 4 \end{bmatrix} \begin{bmatrix} \frac{3}{4}x_1[2n] + \frac{3}{4}x_2[2n] - \frac{3}{4}x_3[2n] \\ x_4[2n] \end{bmatrix}$$
$$+ \begin{bmatrix} \frac{3}{4}(3w_1[2n] + w_1[2n+1]) + \frac{3}{4}(3w_2[2n] + w_2[2n+1]) - \frac{3}{4}(-3w_3[2n] + w_3[2n+1]) \\ 2w_4[2n] + w_4[2n+1] \end{bmatrix}$$

$$\begin{bmatrix} \frac{1}{4} & 0 \end{bmatrix} \mathbf{y}[2n+1] = \beta[2n+1] \left(\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{3}{4}x_1[2n] + \frac{3}{4}x_2[2n] - \frac{3}{4}x_3[2n] \\ x_4[2n] \end{bmatrix} + \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} w_1[2n] \\ w_2[2n] \\ w_3[2n] \\ w_4[2n] \end{bmatrix} \right)$$

where $\frac{3}{4}x_1[2n] + \frac{3}{4}x_2[2n] - \frac{3}{4}x_3[2n]$ and $x_4[2n]$ are the new states of the system.

Here, we can notice that by considering functions of the original states as new states, we get new systems with scalar observations and no eigenvalue cycles. The reduction of the original systems to such systems is what Claim 7.7 does.

Now, we can apply Lemma 7.25 to estimate the states of the new systems, i.e. the estimation of $\frac{1}{4}x_1[2n] - \frac{3}{4}x_2[2n] + \frac{3}{4}x_3[2n]$ and $\frac{3}{4}x_1[2n] + \frac{3}{4}x_2[2n] - \frac{3}{4}x_3[2n]$. Here, we are estimating a function of states instead of the original states themselves. This idea of function decoding was recently proposed and found to be useful in communication problems [73]. Furthermore, the sum of these two functions is $x_1[2n]$, which means we can also estimate $x_1[2n]$ based on the estimation of the functions.

Therefore, we get an estimation of $x_1[n]$, $\hat{x}_1[n]$ (this step corresponds to Claim 7.8) and subtract it from the original system as follows.

$$\mathbf{y}[n] - \beta[n] \begin{bmatrix} 1 \\ 1 \end{bmatrix} \hat{x}_1[n] = \beta[n] \left(\begin{bmatrix} 1 & 1 & 2 \\ -1 & 2 & 3 \end{bmatrix} \begin{bmatrix} x_2[n] \\ x_3[n] \\ x_4[n] \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} x_1[n] - \begin{bmatrix} 1 \\ 1 \end{bmatrix} \hat{x}_1[n] \right)$$

Finally, the resulting system can be thought as of a system with only three states (one state less than the original system). Using induction hypothesis, we can estimate the remaining three states.

2.9 Comments

The intermittent Kalman filtering problem was first motivated by control over communication channels. Therefore, the problem is conventionally believed to fall into the intersection of control and communication. However, if the plant is unstable the transmission power of the sensor diverges to infinity if it is really going to pack an ever increasing number of bits in each transmission. Therefore, it is hard to say that intermittent Kalman filtering has a direct connection to communication theory. Instead, we propose that the intersection of control and signal processing —

especially sampling theory — is the right conceptual category for intermittent Kalman filtering. It should thus be interesting to explore the connection of the results of this chapter with classical and modern results in sampling theory.

Arguably, the closest problem to intermittent Kalman filtering is that of observability after sampling. As we mentioned earlier, the observability of $(\mathbf{A_c}, \mathbf{C_c})$ in (2.19) and (2.20) does not imply the observability of $(\mathbf{A_c}, \mathbf{C})$ in (2.26) and (2.27). The well-known sufficient condition is:

Theorem 2.9 (Theorem 6.9. of [17]). *Suppose $(\mathbf{A_c}, \mathbf{C_c})$ is observable. A sufficient condition for its discretized system with sampling interval I to be observable is that $\frac{|\Im(\lambda_i - \lambda_j)I|}{2\pi} \notin \mathbb{N}$ whenever $\Re(\lambda_i - \lambda_j) = 0$.*

The eigenvalues of the sampled system are given as $\exp(\lambda_i I)$. Thus, the above theorem tells that when the sampling does not map two distinct eigenvalues to the same one, the sampled system is also observable.

For intermittent observability, we can write a similar theorem. When the sampling does not make two distinct eigenvalues belong to the same eigenvalue cycle, the sampled system has the critical erasure probability of $\frac{1}{|e^{2\lambda_{max} I}|}$.

Corollary 2.3. *Suppose $(\mathbf{A_c}, \mathbf{C_c})$ is observable. A sufficient condition for its discretized system with sampling interval I to have $\frac{1}{|e^{2\lambda_{max} I}|}$ as a critical erasure probability is that $\frac{|\Im(\lambda_i - \lambda_j)I|}{2\pi} \notin \mathbb{Q}$ whenever $\Re(\lambda_i - \lambda_j) = 0$.*

Proof immediately follows from Corollary 2.1 and the fact that the eigenvalues of the sampled system are $\exp(\lambda_i I)$.

The idea of breaking cyclic behavior using non-uniform sampling is also shown in the context of sampling multiband signals [79]. The lower bound on the sampling rate is known to be the Lebesgue measure of the spectral support of the signal sampled. To achieve this lower bound for a general multiband signal, a nonuniform sampling pattern has to be used. Moreover, nonuniform sampling is also well known as a necessary condition for the currently hot field of compressed sensing [25].

As a last comment, we would like to mention that the result is not sensitive to the norm. In this chapter, intermittent observability is defined using the l^2-norm to follow the majority of the literature. But, if intermittent observability is defined by the l^η-norm, we can simply replace 2 in every theorem by η. For example, the result of Theorem 2.7 becomes $\frac{1}{\max_i |\lambda_{i,1}|^{\frac{\eta p_i}{l'_i}}}$.

Chapter 3

Network Coding meets Decentralized Control

3.1 Introduction

This chapter is inspired by the similarity between the algebraic characterization of fixed modes [4] in decentralized control problems and the min-cut bound in information theory [21].

Consider a standard decentralized linear system

$$\begin{aligned} x[n+1] &= Ax[n] + B_1 u_1[n] + \cdots + B_v u_v[n] \\ y_1[n] &= C_1 x[n] \\ &\vdots \\ y_v[n] &= C_v x[n]. \end{aligned}$$

Here, the input $u_i[n]$ must be a causal LTI functions of the observations $y_i[n]$. Then, the algebraic condition for λ to be a fixed mode [4, Theorem 4.1] is

$$\min_{V \subseteq \{1,2,\cdots,v\}} \text{rank} \begin{bmatrix} A - \lambda I & B_V \\ C_{V^c} & 0 \end{bmatrix} \geq \dim(A). \tag{3.1}$$

If λ is a fixed mode, that implies that no LTI control strategy can stabilize that mode. Consider a communication relay network shown in [21, Theorem 15.10.1] where the input to the channel at the relay node i is X_i and the output from the channel at the relay node i is Y_i. Then, the information-theoretic min-cut bound [21, Theorem 15.10.1] is

$$\min_{V \subseteq \{1,2,\cdots,v\}} I(X_V; Y_{V^c} | X_{V^c}) \geq \sum_{i \in V, j \in V^c} R_{ij}. \tag{3.2}$$

We can see that the left-hand sides of both (3.1) and (3.2) have a minimization over all subsets V. Moreover, in noiseless relay networks the mutual information is essentially equal to the rank of an appropriate channel matrix[1] [99]. Therefore, the left-hand sides of (3.1) and (3.2) can be considered to be exactly the same. Identifying the right hand sides of (3.1) and (3.2) with each other, we can see that the dimension of A seems to correspond to a rate of total information flow. Moreover, fixed modes are closely connected to stabilizability. Thus, we can conjecture that a decentralized system is stabilizable if and only if enough information flow can be supported to stabilize the plant, and vice versa. In this chapter, we make this conjecture rigorous.

First, let's review perspectives on information flow in communication networks. Historically, information in a network was believed to behave like a physical commodity. The network was modeled using a graph, and the information was thought of as commodities to be transported from the source to the destination by routing them through the nodes. The most important result is the celebrated mincut-maxflow theorem [28, 31], which reveals that the maximum amount of commodity flow through a graph is equal to the minimum cut of the graph. Moreover, this maximum flow is achievable by a routing scheme. For decades, this optimality result made researchers stick to routing solutions even for information.

However, in [1] it was found that information flow in networks does not really behave like physical commodities do. Obviously, we can copy information. But going further, we can also process and mix information. The famous *butterfly example* shows that for multiple-source multiple-destination cases, there is a gain by allowing relays to mix their incoming signals instead of just routing them.

Even if physical commodity flows (which we can only route) and information flows (which we can copy, process and mix) are different, the graph-theoretic concepts and insights originally developed for commodity flows continue to be helpful. The main difference is that the amount of flow, which is naturally measured by the number (or weight or volume) of commodities in physical commodity flows, must instead be measured in "dimensions" of the signal for information flows. However, the mincut-maxflow theorem remains the main tool to understand network information flows. For example, in the multicast problem the relevant mincut is the minimum of the mincut to each destination, and the mincut-maxflow theorem still holds [1]. Moreover, this maximum flow is achievable by linear time-invariant network coding [52].

Once information-theorists had the freedom to mix and process signals inside the nodes that they could design, they also started to consider such operations as potentially existing outside

[1] Information is traditionally measured in bits and the rate of bits that a channel can carry is computed by the mutual information $I(X;Y)$. However, in continuous-alphabet channels like the AWGN (additive white Gaussian noise) channel, the mutual information depends crucially on the signal-to-noise ratio and scales as log SNR. It was noticed that when the channel has multiple-inputs and multiple-outputs (MIMO) — like when there are multiple antennas involved in wireless communication — the mutual information increases as the rank of the channel matrix times log SNR. This fact inspired the creation of the finite-field noiseless MIMO channel model, within which the mutual information is equal to the rank of the channel matrix multiplied by the log of the field size. Therefore, the rank can be considered another measure for information, as measured in units of dimensions or degrees-of-freedom. We refer the reader to [99] for further details.

these nodes [74]. The signals from the relay nodes could be broadcast to multiple receiving nodes or superposed with other signals at a receiving node. In fact, such extensions were a natural fit to wireless communication [6]. The operations outside the nodes modeled communication channels and such wireless channel models had long been valuable even when restricted to be linear time-invariant.

At this point, we can see the similarities between network-coding problems [6] and decentralized-linear-control problems [104]. The network channels (which we cannot design) can be considered as the linear plant. The source, relays and destination nodes (which we can design) can be considered as decentralized controllers. Just as decentralized controllers process and combine their observations to generate their control inputs, the relay nodes process and combine their incoming signals from the channel to generate their outgoing signals.

Despite these similarities, many differences between the communication and control problems had been preventing a firm connection being made between them. First of all, network-coding information-theorists work in finite fields, whereas control-theorists default to infinite fields like the reals or complex numbers. Moreover, information-theorists tend not to have any explicit state in the system, preferring an input-output perspective. Most importantly, the information-theorists have a clearly specified source and destination, and their goal is to push information from one to the other. The control-theorists tend not to have explicit sources and destinations, and instead there is a dynamic evolution that needs to be controlled or stabilized.

The main goal of this chapter is to bridge these differences and make a concrete connection between network coding and decentralized linear control. We first apply linear-system-theoretic ideas to network coding to propose *network linearization* as an algorithm to convert an arbitrary-topology network to an equivalent acyclic single-hop relay network. Based on this, we prove an algebraic mincut-maxflow theorem, Theorem 3.2.

Then, we apply network coding ideas to decentralized linear systems. As shown in Theorem 3.7 and 3.8, we prove that if a decentralized linear system is LTI[2]-stabilizable, then there must exist a corresponding implicit information flow sufficient to stabilize the system.

The rest of the chapter is organized as follows: In Section 3.2, we introduce the definition of LTI networks and prove an algebraic mincut-maxflow theorem based on network linearization. We also compare network linearization with the known idea of network unfolding. In Section 3.3, we introduce some preliminary facts about decentralized linear systems. Section 3.4 shows a representative example that clearly illustrates the implicit information flows in decentralized control systems. Section 3.5 gives the capacity-stabilizability equivalence theorem. In Section 3.6, we consider the stabilizability problem with an explicit communication network, and convert networking results to the equivalent stabilizability results.

[2]It is in our focus on stabilizability using only linear *time-invariant* control laws that the results in this chapter differ from the results in [116] where *time-varying* control laws are permitted. The overall perspectives however are compatible in that we are also interested in cutsets and information flows.

3.2 LTI Communication Networks

3.2.1 Definitions and Algebraic Mincut-Maxflow Theorem

An LTI communication network is a collection of transmitters, relays, and receivers — which will be called nodes.[3]

Each node has input and output ports. These connect to the channels. Each node generates a signal and sends it to the channels through its output ports, which are simultaneously the input to the channels. In this chapter, we model signals elements from a field $\mathbb{F}$ and time is discrete. The transmitted signals go through the channels and arrive at the channel outputs, which are simultaneously the input ports of the nodes. We take a channel-centric perspective in this chapter's notation.

The relationship between the input and output signals of the channels is given by nature. In LTI communication networks, the input-output relationships of the channels are linear time-invariant. Thus, they can be described by transfer functions. Furthermore, since we will focus on discrete-time systems, by taking z-transforms the transfer functions can be represented by rational functions in z.

Even though the channels are given by nature, we still have design freedom for the nodes. Each node can choose the input signals to the channels as arbitrary causal functions on the output signals from the channels. In LTI networks, the node operation is restricted to be linear time-invariant. In other words, the nodes can be thought as causal linear time-invariant filters between the output signals from the channels into the input signals to the channels. To reflect this design freedom, we will assign different variables k_i for the transfer functions inside the nodes.

We focus on LTI point-to-point communication networks with one transmitter and one receiver, and we denote the network as $\mathcal{N}(z)$. Let's formally define LTI point-to-point networks using graph notation. The input and output ports of the nodes can be modeled as vertices. The transfer functions connecting them can be thought as directed edges. Consider a digraph (W, E) with a totally ordered set of vertices (ports) W and a set of edges E. W is partitioned according to which node that port belongs to.

In other words, for an LTI network with v relays, W can be partitioned into the sets N_{tx}, $N_1, \cdots, N_v, N_{rx}$, i.e. $N_i \subseteq W$, $N_i \cap N_j = \emptyset$ for $i \neq j$, and $\bigcup_{i \in \{tx,1,\cdots,v,rx\}} N_i = W$. Thus, a set of vertices N_i corresponds to a node.

To simplify the notation, we will use the subscript "tx" and -1 interchangeably. Likewise, we will also use the subscript $v+1$ for the subscript "rx", i.e. $N_{tx} = N_0$ and $N_{rx} = N_{v+1}$.

For a given node N_i, the elements of N_i are again partitioned into two subsets $N_{i,in}$ and $N_{i,out}$ which are called the input and output vertices of the node i. The inputs and the outputs are

[3] The LTI networks considered here are essentially the same as the linear deterministic model studied in [6] except that our LTI networks restrict the relay design to be linear time-invariant and the underlying field can also be real or complex instead of being restricted to finite fields.

defined in a channel-centric perspective. So an input vertex is an output port of a node, and an output vertex is an input port of the node. $N_{i,in}$ represent the signals going out from the node i *into* the channels and $N_{i,out}$ represent the signals coming *out* from the channels into the node i.

The transmitter node does not receive signals and the receiver node does not transmit signals, so $N_{tx,out} = \emptyset$ and $N_{rx,in} = \emptyset$. We denote the number of the input and output vertices of the node i as $d_{i,in}$ and $d_{i,out}$, i.e. $d_{i,in} := |N_{i,in}|$, and $d_{i,out} := |N_{i,out}|$.

Let the signals take values from a field $\mathbb{F}$, let z be the dummy variable for z-transforms, and let $K = \{k_1, k_2, k_3, \cdots\}$ be a set of variables to represent the gains inside the nodes. We also define $\mathbb{F}[z]$, $\mathbb{F}[K]$, $\mathbb{F}[z, K]$ as the field of all rational functions in variables z, K, $\{z\} \cup K$ with coefficients in $\mathbb{F}$ respectively.

Each edge which connects the ports of the nodes can be written as a triplet $(w', w'', h_{w',w''}(z, K))$ $\in E$ where $w', w'' \in W$ and $h_{w',w''}(z, K) \in \mathbb{F}[z] \cup K$. Here, w' is the starting port of the edge, w'' is called the end of the edge, and $h_{w',w''}(z, K)$ is the gain of the connection.

Since a lack of physical connection between two vertices w' and w'' can be represented as $h_{w',w''}(z, K) = 0$, we assume that every input vertex is connected to every output vertex, including "self-loops" connecting the input vertices to its own output vertices. There are two kinds of edges. One kind of edges is the transfer functions connecting the input vertices to the output vertices —channel transfer functions. They are given by nature and described by z-transforms —rational functions on z. Formally, for all $i, j \in \{0, \cdots, v + 1\}$ and $w' \in N_{in,i}, w'' \in N_{out,j}$,

$$(w', w'', h_{w',w''}(z, K)) \in E \text{ and } h_{w',w''}(z, K) \in \mathbb{F}[z].$$

The other kind of edge is inside each node. There we have design freedom. To reflect this, for each node let there exist edges fully connecting its output vertices to its input vertices. The transfer functions associated with these edges are in the form of $k_i \in K$ and distinct. Since the transmitter and receiver have only one kind of ports, N_{tx} and N_{rx} do not have internal edges.

This distinct transfer function assumption guarantees enough design freedom at the relays since we can assign different transfer functions to different edges. Formally, for all $i \in \{1, \cdots, v\}$ and $w' \in N_{out,i}, w'' \in N_{in,i}$, $(w', w'', h_{w',w''}(z, K)) \in E$ and $h_{w',w''}(z, K) = k_{w',w''}$ where $k_{w',w''} \in K$. If (w'_1, w''_1) and (w'_2, w''_2) are distinct internal edges, $h_{w'_1,w''_1} \neq h_{w'_2,w''_2}$. These internal edges represent the potential LTI communication schemes. In a fully realized network with a specific communication scheme, each element of the K will be replaced with a specific element in $\mathbb{F}[z]$.

At each vertex and edge, the signal is processed as follows: Each vertex $w \in W$ adds all the signals coming from the edges whose head is w and transmits to the edges whose tail is w. Each edge $e \in E$ multiplies the signal coming from its tail with its transfer function and transmits this to its head.

Denote a transfer function matrix from the input vertices of the node N_i to the output vertices of the node N_j as $H_{i,j}(z)$. In the same way, we denote a transfer function from a set (ordered set) of nodes A to a set (ordered set) of nodes B as $H_{A,B}(z)$. We also denote the transfer function matrix from the output vertices (input ports) of N_i to the input vertices (output vertices) of N_i as K_i. Then, $H_{i,j}(z) \in \mathbb{F}[z]^{d_{j,out} \times d_{i,in}}$ and $K_i \in \mathbb{F}[K]^{d_{i,in} \times d_{i,out}}$. For briefness, we write $H_{i,j}(z)$ as $H_{i,j}$ when it does not cause confusion. K_i are given in forms of $\begin{bmatrix} k_{i_1} & k_{i_2} & \cdots \\ k_{i_3} & k_{i_4} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$.

As mentioned above, by considering the transfer functions of the internal edges as different bare dummy variables in K, we reflect the design freedom of the relay nodes. Moreover, the capacity of a network —the rank of the transfer function matrix— will be maximized by considering the transfer functions of the internal edges as variables in K. Precisely, let $K_i(z) \in \mathbb{F}[z]^{d_{i,in} \times d_{i,out}}$ be a matrix whose size is the same as K_i but the elements of the matrix belong to $\mathbb{F}[z]$. Denote the transfer functions from the transmitter to the receiver of $\mathcal{N}(z)$ as $G(z, K)$ and $G(z, K(z))$ in each case. Then, we have the following relationship:

Lemma 3.1. *Let $G(z, K)$ be given as above. Then, we have the following relationship between the rank of $G(z, K)$ and $G(z, K(z))$.*

$$\operatorname{rank} G(z, K) = \max_{K_i(z) \in \mathbb{F}[z]^{d_{i,in} \times d_{i,out}}} \operatorname{rank} G(z, K(z)).$$

Proof. The proof is essentially the same as [52, Lemma 1]. For all $K_i(z) \in \mathbb{F}[z]^{d_{i,in} \times d_{i,out}}$ the independent columns in $G(z, K(z))$ are still independent even if we consider the elements of K_i as variables. Therefore, for all $K_i(z) \in \mathbb{F}[z]^{d_{i,in} \times d_{i,out}}$, $\operatorname{rank} G(z, K) \geq \operatorname{rank} G(z, K(z))$.

Moreover, the rational function field $\mathbb{F}[z]$ has an infinite number of elements and the dimension of the algebraic variety that makes $G(z, K)$ lose its rank is strictly smaller than the dimension of K_i's. Therefore, there exists $K_i(z) \in \mathbb{F}[z]^{d_{i,in} \times d_{i,out}}$ such that $\operatorname{rank} G(z, K) = \operatorname{rank} G(z, K(z))$. Thus, the lemma is true. □

Figure 3.1 shows the graphical representation of an LTI communication network. The squares represent the nodes of the LTI networks. The empty circles attached to the squares represent the input vertices (output ports) from the nodes to the channels. The circles with plus represent the output vertices (input ports) from the channels to the nodes. The arrows outside the nodes (connecting empty circles to plus circles) represent the communication channels, and the arrows inside the nodes (connecting plus circles to empty circles) represent the communication schemes. The scalars (or matrices) written on the arrows represent the transfer functions (or transfer function matrices). We also denote a $m \times m$ identity matrix as I_m.

Let $G(z, K)$ be the transfer function from the input vertices of the transmitter node to the output vertices of the receiver node. $G(z, K)$ can be written in terms of $H_{i,j}$ and K_i [75].

Theorem 3.1. *With the above definitions, the transfer function matrix $G(z, K)$ is given as*

$$G(z,K) = \begin{bmatrix} H_{1,rx}K_1 & \cdots & H_{v,rx}K_v \end{bmatrix}$$
$$\left(I - \begin{bmatrix} H_{1,1}K_1 & \cdots & H_{v,1}K_v \\ \vdots & \ddots & \vdots \\ H_{1,v}K_1 & \cdots & H_{v,v}K_v \end{bmatrix} \right)^{-1} \begin{bmatrix} H_{tx,1} \\ \vdots \\ H_{tx,v} \end{bmatrix} + H_{tx,rx}.$$

Proof. As illustrated in Fig. 3.1, let U, X_i and Y be vectors of signals at the input vertices of the transmitter, the output vertices visible at node i, and the output vertices visible at the receiver. Then, we have the following relations between U, X_i and Y:

$$\begin{bmatrix} X_1 \\ \vdots \\ X_v \end{bmatrix} = \begin{bmatrix} H_{1,1}K_1 & \cdots & H_{v,1}K_v \\ \vdots & \ddots & \vdots \\ H_{1,v}K_1 & \cdots & H_{v,v}K_v \end{bmatrix} \begin{bmatrix} X_1 \\ \vdots \\ X_v \end{bmatrix} + \begin{bmatrix} H_{tx,1} \\ \vdots \\ H_{tx,v} \end{bmatrix} U$$

$$Y = \begin{bmatrix} H_{1,rx}K_1 & \cdots & H_{v,rx}K_v \end{bmatrix} \begin{bmatrix} X_1 \\ \vdots \\ X_v \end{bmatrix} + H_{tx,rx}U.$$

Simple algebra then gives the theorem. Here, the invertibility of the matrix can be shown as follows: As shown in Lemma 3.1, the rank of $(I - \begin{bmatrix} H_{1,1}K_1 & \cdots & H_{v,1}K_v \\ \vdots & \ddots & \vdots \\ H_{1,v}K_1 & \cdots & H_{v,v}K_v \end{bmatrix})$ is the largest rank over all $K_i(z)$. Furthermore, by putting $K_i(z) = 0$, the matrix becomes invertible. □

Therefore, from an end-to-end perspective, the point-to-point LTI network $\mathcal{N}(z)$ can be thought as a MIMO (multiple-input multiple-output) channel whose channel matrix is $G(z, K)$. It is well-known that the capacity of MIMO channels is closely related to the rank of the channel matrix [99].

Definition 3.1 (Degree of Freedom Capacity)**.** *For a given LTI network $\mathcal{N}(z)$, we say that the degree of freedom (d.o.f.) capacity of the network $\mathcal{N}(z)$ is k if its transfer matrix $G(z, K)$ is rank k, i.e.* $\text{rank}(G(z, K)) = k$.

On the other hand, when we "cut" the nodes into two disjoint sets $V = \{tx, i_1, \cdots, i_k\}$ and $V^c = \{rx, i_{k+1}, \cdots, i_v\}$, the channel matrix between these two is defined as

$$H_{V,V^c} = \begin{bmatrix} H_{tx,rx} & H_{i_1,rx} & \cdots & H_{i_k,rx} \\ H_{tx,i_{k+1}} & H_{i_1,i_{k+1}} & \cdots & H_{i_k,i_{k+1}} \\ \vdots & \vdots & \ddots & \vdots \\ H_{tx,i_v} & H_{i_1,i_v} & \cdots & H_{i_k,i_v} \end{bmatrix}.$$

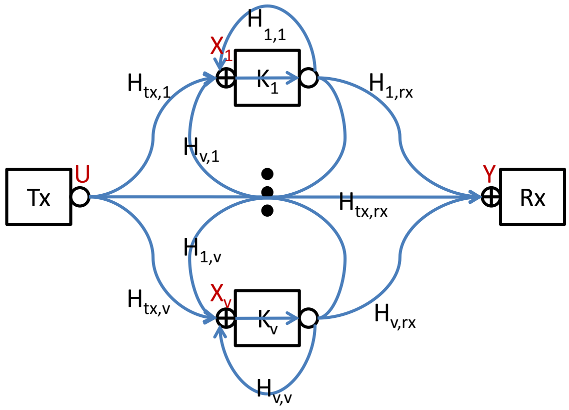


Figure 3.1: point-to-point LTI network $\mathcal{N}(Z)$

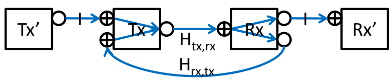


Figure 3.2: We can model feedback by introducing an outer transmitter Tx' and receiver Rx'

Definition 3.2 (Degree of Freedom Mincut)**.** *For a given LTI network $\mathcal{N}(z)$, we say that the degree of freedom (d.o.f.) mincut of the network $\mathcal{N}(z)$ is k if the minimum rank of cuts is equal to k, i.e.* $\min_{V:V\subseteq\{0,\cdots,v+1\},V\ni tx,V\not\ni rx} \operatorname{rank} H_{V,V^c}(z) = k$.

One key fact about LTI networks is that the well-known mincut-maxflow theorem [31, 28] can be extended to them. This is one of the main theorems of the chapter.

Theorem 3.2 (Algebraic Mincut-Maxflow Theorem)**.** *With the above definitions,*

$$\begin{aligned} &\operatorname{rank} G(z,K) \\ &= \min_{V:V\subseteq\{0,\cdots,v+1\},V\ni tx,V\not\ni rx} \operatorname{rank} H_{V,V^c}(z). \end{aligned}$$

Proof. See Section 3.2.2. □

In this theorem, K_i are considered as dummy variables which are independent from z and each other. However, what this theorem really implies is the existence of mincut-achieving coding schemes, i.e. there exist z-transforms that we can plug in for K_i without changing the equality of Theorem 3.2. In Section 3.2.3, we will discuss this point in further detail.

The above notations for LTI point-to-point networks can be naturally generalized to those for LTI networks with multiple sources and destinations.

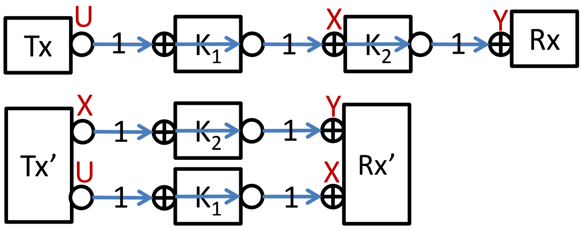


Figure 3.3: LTI network example and its equivalent network with linearized transfer function

One may think the LTI networks above do not cover channels with feedback since we did not include any channel from the receiver to the transmitter. However, as shown in Fig. 3.2 the channel with feedback can be modeled by introducing an outer transmitter and receiver. In a similar way, we can also include cooperation between transmitters and receivers in cases with multiple sources and destinations.

3.2.2 State-Space Representation and Network Linearization

In this section, we prove Theorem 3.2 using the idea of network linearization. Network linearization is the counterpart of the following fact of linear system theory: Every causal linear time-invariant system with an input $u[n]$ and an output $y[n]$ can be written in state-space form [17], *i.e.* can be realized as a linear system equation:

$$x[n+1] = Ax[n] + Bu[n]$$
$$y[n] = Cx[n] + Du[n]$$

by introducing the proper internal states $x[n]$. Similarly, network linearization tells us that every LTI network with an arbitrary topology can be converted to an acyclic single-hop relay network by introducing proper internal states.

First, we illustrate two key ideas for network linearization.

(1) Internal States: Consider the two-hop relay network shown in the top figure of Fig. 3.3. The transfer function from U to Y is k_2k_1, which is not linear in k_1, k_2. To write the transfer function in a linear form, we introduce an internal state X at the output of the second node. Then, the transfer function matrix from X, U to Y, X is $\begin{bmatrix} Y \\ X \end{bmatrix} = \begin{bmatrix} k_2 & 0 \\ 0 & k_1 \end{bmatrix} \begin{bmatrix} X \\ U \end{bmatrix}$, which is linear in k_1, k_2. Moreover, since

$$\begin{bmatrix} k_2 & 0 \\ 0 & k_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} k_1 \begin{bmatrix} 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} k_2 \begin{bmatrix} 1 & 0 \end{bmatrix},$$

it corresponds to the transfer function of the acyclic single-hop relay network shown in the bottom figure of Fig. 3.3.

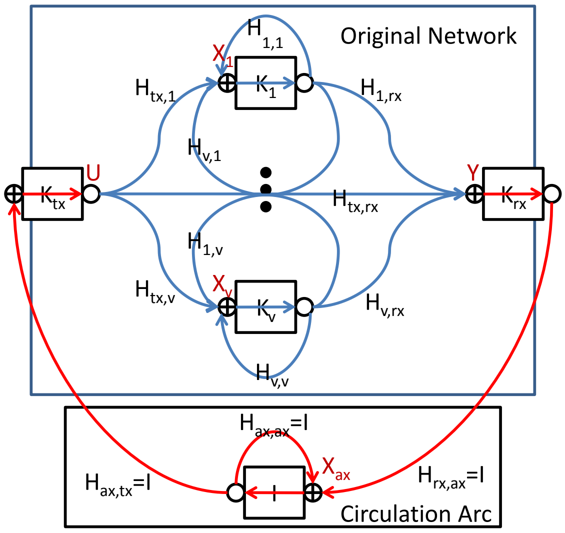


Figure 3.4: LTI network $\mathcal{N}(Z)$ with a circulation arc added in.

(2) Circulation Arc: Even if the transfer function can be written in a linear matrix form by introducing internal states, there has to be a relationship between the rank of the original transfer function and the rank of the linearized transfer function.

After all, in general the rank of the linearized transfer function matrix will be bigger as the above example illustrates. So we need a way to relate the ranks of the transfer function matrices.

To make this connection, we borrow the circulation arc idea from the integer programming context [46, p.86]. The problem that they had was that when they tried to write the maxflow problem in linear programming form, the flow conservation law did not hold at the source and the destination. The flow at the source is negative and the flow at the destination is positive. To patch this, they introduced a circulation arc with infinite capacity from the destination to the source. Since the amount of the negative flow at the source is the same as the amount of the positive flow at the destination, the flow conservative law can be recovered as a universal. Moreover, the flow across the network can be easily measured by measuring the flow in the circulation arc.

To apply this idea to LTI networks, we use an underdetermined system. Let's consider $x = x + K_{rx}G(z,K)K_{tx}x$ with unknown vector x. Here, $K_{rx}G(z,K)K_{tx}$ is a transfer function with a preprocessing matrix K_{tx} and a postprocessing matrix K_{rx}. If the rank of $K_{rx}G(z,K)K_{tx}$ is smaller than the dimension of x, the equation is underdetermined. Otherwise, it is not. Thus, we can see that the rank of the transfer function can be measured by the underdeterminedness of the system.

Now, we will combine these ideas for network linearization. We first formally introduce the circulation arc. As shown in Fig. 3.4, an auxiliary node N_{ax} with d_{ax} input ports and d_{ax} output ports is added to the original network. We also introduce d_{ax} input vertices at the receiver node and d_{ax} output vertices at the transmitter node. Let $H_{rx,ax} = H_{ax,tx} = H_{ax,ax} = K_{ax} = I_{d_{ax}}$. As discussed in Section 3.2.1, to reflect the design freedom of the transmitter and receiver, let $K_{tx} \in \mathbb{F}[K]^{d_{tx} \times d_{ax}}$ and $K_{rx} \in \mathbb{F}[K]^{d_{ax} \times d_{rx}}$, and each element of K_{tx}, K_{rx} is the form of $k_i \in K$ and they are all distinct and also distinct from the elements in $K_1, \cdots, K_v$ inside the relays.

Now, we introduce labels for the internal states. As shown in Fig. 3.4, let X_{ax}, X_i, and Y be the vectors of the signals of the output vertices seen at the auxiliary node, the node i, and the receiver respectively.

From the system diagram, Fig. 3.4, we can see the following relation has to hold.

$$\begin{bmatrix} X_{ax} \\ Y \\ X_1 \\ \vdots \\ X_v \end{bmatrix} = \begin{bmatrix} I_{d_{ax}} & K_{rx} & 0 & \cdots & 0 \\ H_{tx,rx}K_{tx} & 0 & H_{1,rx}K_1 & \cdots & H_{v,rx}K_v \\ H_{tx,1}K_{tx} & 0 & H_{1,1}K_1 & \cdots & H_{v,1}K_v \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ H_{tx,v}K_{tx} & 0 & H_{1,v}K_1 & \cdots & H_{v,v}K_v \end{bmatrix} \begin{bmatrix} X_{ax} \\ Y \\ X_1 \\ \vdots \\ X_v \end{bmatrix}$$

$$(\Leftrightarrow) \underbrace{\begin{bmatrix} 0 & -K_{rx} & 0 & \cdots & 0 \\ -H_{tx,rx}K_{tx} & I_{d_{rx}} & -H_{1,rx}K_1 & \cdots & -H_{v,rx}K_v \\ -H_{tx,1}K_{tx} & 0 & I_{d_{1,out}} - H_{1,1}K_1 & \cdots & -H_{v,1}K_v \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -H_{tx,v}K_{tx} & 0 & -H_{1,v}K_1 & \cdots & I_{d_{v,out}} - H_{v,v}K_v \end{bmatrix}}_{:=G_{lin}(z,K)} \begin{bmatrix} X_{ax} \\ Y \\ X_1 \\ \vdots \\ X_v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (3.3)$$

The matrix $G_{lin}(z,K)$ here is filled with entries linear in K_i. Thus, $G_{lin}(z,K)$ can be rewritten as

$$G_{lin}(z,K) = \underbrace{\begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & I & 0 & \cdots & 0 \\ 0 & 0 & I & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & I \end{bmatrix}}_{:=A} + \underbrace{\begin{bmatrix} 0 \\ H_{tx,rx} \\ H_{tx,1} \\ \vdots \\ H_{tx,v} \end{bmatrix}}_{:=B_{tx}} K_{tx} \underbrace{\begin{bmatrix} -I_{d_{ax}} & 0 & 0 & \cdots & 0 \end{bmatrix}}_{:=C_{tx}}$$

$$+ \underbrace{\begin{bmatrix} 0 \\ H_{1,rx} \\ H_{1,1} \\ \vdots \\ H_{1,v} \end{bmatrix}}_{:=B_1} K_1 \underbrace{\begin{bmatrix} 0 & 0 & -I_{d_{1,out}} & \cdots & 0 \end{bmatrix}}_{:=C_1} + \cdots + \underbrace{\begin{bmatrix} 0 \\ H_{v,rx} \\ H_{v,1} \\ \vdots \\ H_{v,v} \end{bmatrix}}_{:=B_v} K_v \underbrace{\begin{bmatrix} 0 & 0 & 0 & \cdots & -I_{d_{v,out}} \end{bmatrix}}_{:=C_v}$$

$$+ \underbrace{\begin{bmatrix} I_{d_{ax}} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{:=B_{rx}} K_{rx} \underbrace{\begin{bmatrix} 0 & -I_{d_{rx}} & 0 & \cdots & 0 \end{bmatrix}}_{:=C_{rx}}. \tag{3.4}$$

The $A, B_{tx}, C_{tx}, B_i, C_i, B_{rx}, C_{rx}$ are defined as above in (3.4).

Because $G_{lin}(z,K)$ looks like a transfer function matrix, we can formally ask what is the LTI network whose transfer function matrix is $G_{lin}(z,K)$. Then, we can easily see that $G_{lin}(z,K)$ corresponds to the transfer function of the linearized LTI network $\mathcal{N}_{lin}(z)$ of Fig. 3.5. The linearized network $\mathcal{N}_{lin}(z)$ has a new transmitter tx' and receiver rx', and is an acyclic single-hop relay network with a direct link between tx' and rx'. We also use the subscript "tx'" and -1 alternatively, and likewise "rx'" and $v+2$ alternatively.

Let $d := \dim \begin{bmatrix} Y \\ X_1 \\ \vdots \\ X_v \end{bmatrix} = d_{rx} + \sum_{1 \le i \le v} d_{i,out}$ where $Y, X_1, \cdots, X_v$ are given as (3.3). Then, we will prove that the maxflow of $\mathcal{N}_{lin}(z)$ is the same as the maxflow of $\mathcal{N}(z)$ by an offset d.

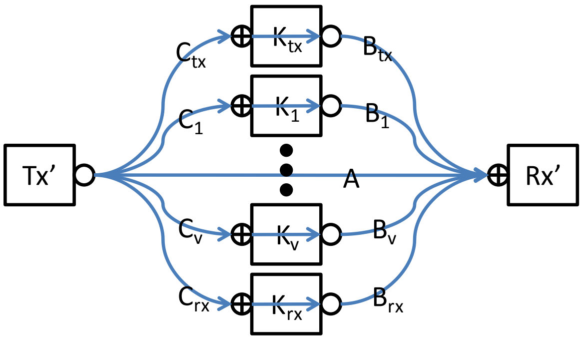


Figure 3.5: Linearized LTI network $\mathcal{N}_{lin}(z)$

Furthermore, for sets (ordered sets) $V = \{v_1, \cdots, v_i\}$ and $W = \{w_1, \cdots, w_j\}$ we define

$$B_V := \begin{bmatrix} B_{v_1} & \cdots & B_{v_i} \end{bmatrix}$$

$$C_V := \begin{bmatrix} C_{v_1} \\ \vdots \\ C_{v_i} \end{bmatrix}$$

$$D_{V,W} := \begin{bmatrix} D_{v_1 w_1} & \cdots & D_{v_1 w_j} \\ \vdots & \ddots & \vdots \\ D_{v_i w_1} & \cdots & D_{v_i w_j} \end{bmatrix}$$

whenever this shorthand does not cause confusion.

We also denote the channel matrices from the node i to the node j in the linearized LTI network $\mathcal{N}_{lin}(z)$ as $H^{lin}_{i,j}$. Then, we can easily see that the channel matrix for the cut $V \subseteq \{0, \cdots, v+1\}$ is

$$H^{lin}_{V \cup \{tx'\}, V^c \cup \{rx'\}} = \begin{bmatrix} A & B_V \\ C_{V^c} & 0 \end{bmatrix}. \tag{3.5}$$

We will prove the essential equivalence between the original network $\mathcal{N}(z)$ and the linearized network $\mathcal{N}_{lin}(z)$. First, we prove a lemma on matrix rank.

Lemma 3.2. *For a field $\mathbb{F}$ and $n_1, n_2 \in \mathbb{Z}^+$, let $A \in \mathbb{F}^{n_1 \times n_1}, B \in \mathbb{F}^{n_2 \times n_1}, C \in \mathbb{F}^{n_1 \times n_2}, D \in \mathbb{F}^{n_2 \times n_2}$. If D is invertible, the following rank equality holds.*

$$\text{rank} \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \text{rank}\, D + \text{rank}(A - BD^{-1}C)$$

Proof.

$$\begin{aligned}
\operatorname{rank}\begin{bmatrix} A & B \\ C & D \end{bmatrix} &= \operatorname{rank}\left(\begin{bmatrix} I_{n_1} & -BD^{-1} \\ 0 & I_{n_2} \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix}\right) \\
&= \operatorname{rank}\begin{bmatrix} A - BD^{-1}C & 0 \\ C & D \end{bmatrix} \\
&= \operatorname{rank} D + \operatorname{rank}(A - BD^{-1}C)
\end{aligned}$$

where the first equality comes from the fact that $\begin{bmatrix} I_{n_1} & -BD^{-1} \\ 0 & I_{n_2} \end{bmatrix}$ is invertible, and the last equality is a consequence of D being invertible. □

Now, we prove that the maxflow of the two networks $\mathcal{N}(z)$ and $\mathcal{N}_{lin}(z)$ are equivalent with an offset d.

Lemma 3.3 (Maxflow Equivalence Lemma)**.** *Given the above notations,*

$$\operatorname{rank}(K_{rx}G(z,K)K_{tx}) + d = \operatorname{rank} G_{lin}(z,K).$$

Proof.

$$
\begin{aligned}
&\operatorname{rank} G_{lin}(z,K) \\
&\stackrel{(A)}{=} \operatorname{rank}\begin{bmatrix} I & -H_{1,rx}K_1 & \cdots & -H_{v,rx}K_v \\ 0 & I-H_{1,1}K_1 & \cdots & -H_{v,1}K_v \\ \vdots & \vdots & \ddots & \vdots \\ 0 & -H_{1,v}K_1 & \cdots & I-H_{v,v}K_v \end{bmatrix} \\
&+\operatorname{rank}\left(-\begin{bmatrix} -K_{rx} & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} I & -H_{1,rx}K_1 & \cdots & -H_{v,rx}K_v \\ 0 & I-H_{1,1}K_1 & \cdots & -H_{v,1}K_v \\ \vdots & \vdots & \ddots & \vdots \\ 0 & -H_{1,v}K_1 & \cdots & I-H_{v,v}K_v \end{bmatrix}^{-1} \begin{bmatrix} -H_{tx,rx}K_{tx} \\ -H_{tx,1}K_{tx} \\ \vdots \\ -H_{tx,v}K_{tx} \end{bmatrix} \right) \\
&\stackrel{(B)}{=} d+\operatorname{rank}\left(-\begin{bmatrix} -K_{rx} & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} I & -H_{1,rx}K_1 & \cdots & -H_{v,rx}K_v \\ 0 & I-H_{1,1}K_1 & \cdots & -H_{v,1}K_v \\ \vdots & \vdots & \ddots & \vdots \\ 0 & -H_{1,v}K_1 & \cdots & I-H_{v,v}K_v \end{bmatrix}^{-1} \begin{bmatrix} -H_{tx,rx}K_{tx} \\ -H_{tx,1}K_{tx} \\ \vdots \\ -H_{tx,v}K_{tx} \end{bmatrix} \right) \\
&\stackrel{(C)}{=} d+\operatorname{rank}\left(K_{rx} \begin{bmatrix} I & \begin{bmatrix} H_{1,rx}K_1 & \cdots & H_{v,rx}K_v \end{bmatrix} \begin{bmatrix} I-H_{1,1}K_1 & \cdots & -H_{v,1}K_v \\ \vdots & \ddots & \vdots \\ -H_{1,v}K_1 & \cdots & I-H_{v,v}K_v \end{bmatrix}^{-1} \end{bmatrix} \begin{bmatrix} -H_{tx,rx}K_{tx} \\ -H_{tx,1}K_{tx} \\ \vdots \\ -H_{tx,v}K_{tx} \end{bmatrix} \right) \\
&\stackrel{(D)}{=} d+ \\
&\operatorname{rank}\left(K_{rx} \left(H_{tx,rx} + \begin{bmatrix} H_{1,rx}K_1 & \cdots & H_{v,rx}K_v \end{bmatrix} \begin{bmatrix} I-H_{1,1}K_1 & \cdots & -H_{v,1}K_v \\ \vdots & \ddots & \vdots \\ -H_{1,v}K_1 & \cdots & I-H_{v,v}K_v \end{bmatrix}^{-1} \begin{bmatrix} H_{tx,1} \\ \vdots \\ H_{tx,v} \end{bmatrix} \right) K_{tx} \right) \\
&\stackrel{(E)}{=} d+\operatorname{rank}(K_{rx}G(z,K)K_{tx})
\end{aligned}
$$

(A): This comes from Lemma 3.2 by considering 0_{d_rx} as A, $\begin{bmatrix} -K_{rx} & 0 & \cdots & 0 \end{bmatrix}$ as B, $\begin{bmatrix} -H_{tx,rx}K_{tx} \\ -H_{tx,1}K_{tx} \\ \vdots \\ -H_{tx,v}K_{tx} \end{bmatrix}$ as C, and $\begin{bmatrix} I_{d_{rx}} & -H_{1,rx}K_1 & \cdots & -H_{v,rx}K_v \\ 0 & I_{d_{1,out}}-H_{1,1}K_1 & \cdots & -H_{v,1}K_v \\ \vdots & \vdots & \ddots & \vdots \\ 0 & -H_{1,v}K_1 & \cdots & I_{d_{v,out}}-H_{v,v}K_v \end{bmatrix}$ as D. Here, D is invertible, since by

Lemma 3.1 the rank of D is the maximum rank over all $K_i(z)$ and by setting $K_i(z) = 0$ the matrix D becomes full rank.

(B): Since each element of K_i is a dummy variable,

$$\text{rank}\begin{bmatrix} I_{d_{rx}} & -H_{1,rx}K_1 & \cdots & -H_{v,rx}K_v \\ 0 & I_{d_{1,out}} - H_{1,1}K_1 & \cdots & -H_{v,1}K_v \\ \vdots & \vdots & \ddots & \vdots \\ 0 & -H_{1,v}K_1 & \cdots & I_{d_{v,out}} - H_{v,v}K_v \end{bmatrix} \geq \text{rank}\begin{bmatrix} I_{d_{ax}} & 0 & \cdots & 0 \\ 0 & I_{d_{1,out}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & I_{d_{v,out}} \end{bmatrix} = d.$$

Moreover, because the dimension of the matrix is $d \times d$, the rank is also upper bounded by d.

(C): We can easily show $\begin{bmatrix} I & B \\ 0 & D \end{bmatrix}^{-1} = \begin{bmatrix} I & -BD^{-1} \\ 0 & D^{-1} \end{bmatrix}$. Thus, by considering $\begin{bmatrix} -H_{1,rx}K_{rx} & \cdots & -H_{v,rx}K_v \end{bmatrix}$ as B and $\begin{bmatrix} I_{d_{rx}} & -H_{1,rx}K_1 & \cdots & -H_{v,rx}K_v \\ 0 & I_{d_{1,out}} - H_{1,1}K_1 & \cdots & -H_{v,1}K_v \\ \vdots & \vdots & \ddots & \vdots \\ 0 & -H_{1,v}K_1 & \cdots & I_{d_{v,out}} - H_{v,v}K_v \end{bmatrix}$ as D, and multiplying with the matrix $\begin{bmatrix} -K_{rx} & 0 & \cdots & 0 \end{bmatrix}$, we can prove this step.

(D): This comes from direct computation.

(E): This comes from the definition of $G(z, K)$ shown in Theorem 3.1. □

The mincut of $\mathcal{N}_{lin}(z)$ is also the same as the mincut of $\mathcal{N}(z)$, except for an offset d.

Lemma 3.4 (Mincut Equivalence Lemma)**.** *Given the above notation,*

$$\begin{aligned} &\min\{\text{rank}\, K_{tx}, \text{rank}\, K_{rx}, \min_{W \subseteq \{0,\cdots,v+1\}, W \ni tx, W \not\ni rx} \text{rank}\, H_{W,W^c}\} + d \\ &= \min_{V \subseteq \{-1,\cdots,v+2\}, V \ni tx', V \not\ni rx'} \text{rank}\, H^{lin}_{V,V^c}. \end{aligned} \tag{3.6}$$

Proof. As we can see in the R.H.S. of (3.6), V is a cut of $\mathcal{N}_{lin}(z)$. We will divide V into three cases: (i) When $tx \in V^c$, (ii) When $rx \in V$, and (iii) When $tx \in V$ and $rx \in V^c$.

For cases (i) and (ii), we will show that the rank of channel matrices is at least $\dim X_{ax} + d$. For case (iii), we will show a one-to-one mapping between the cut V for $\mathcal{N}_{lin}(z)$ and the cut W for $\mathcal{N}(z)$ — essentially V is a cut of the original network $\mathcal{N}(z)$.

(i) When $tx \in V^c$,

Notice that by definition, we have

$$\operatorname{rank}\begin{bmatrix} A \\ C_{tx} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & I_{d_{rx}} & 0 & \cdots & 0 \\ 0 & 0 & I_{d_{1,out}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & I_{d_{v,out}} \\ -I_{d_{ax}} & 0 & 0 & \cdots & 0 \end{bmatrix} = \dim X_{ax} + d.$$

Moreover, whenever $tx \in V^c$, the channel matrix for the cut H^{lin}_{V,V^c} contains $\begin{bmatrix} A \\ C_{tx} \end{bmatrix}$ and so $\operatorname{rank} H^{lin}_{V,V^c} \geq \dim X_{ax} + d$. Thus, we have

$$\min_{V \subseteq \{-1,\cdots,v+2\}, V \ni tx', V \not\ni rx', V^c \ni tx} \operatorname{rank} H^{lin}_{V,V^c} \geq \dim X_{ax} + d. \tag{3.7}$$

Furthermore, by choosing $V = \{tx'\}$, we have

$$\operatorname{rank} H^{lin}_{tx',\{tx,1,\cdots,v,rx,rx'\}} = \operatorname{rank}\begin{bmatrix} A \\ C_{tx} \\ C_1 \\ \vdots \\ C_v \\ C_{rx} \end{bmatrix} = \dim X_{ax} + d. \tag{3.8}$$

Therefore, by (3.7) and (3.8) we can conclude

$$\min_{V \subseteq \{-1,\cdots,v+2\}, V \ni tx', V \not\ni rx', V^c \ni tx} \operatorname{rank} H^{lin}_{V,V^c} = \dim X_{ax} + d.$$

(ii) When $rx \in V$,

Notice that by definition, we have

$$\operatorname{rank}\begin{bmatrix} A & B_{rx} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & I_{d_{ax}} \\ 0 & I_{d_{rx}} & 0 & \cdots & 0 & 0 \\ 0 & 0 & I_{d_{1,out}} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & I_{d_{v,out}} & 0 \end{bmatrix} = \dim X_{ax} + d.$$

Moreover, whenever $rx \in V$, the channel matrix for the cut H^{lin}_{V,V^c} contains $\begin{bmatrix} A & B_{rx} \end{bmatrix}$ and so $\operatorname{rank} H^{lin}_{V,V^c} \geq \dim X_{ax} + d$. Thus, we have

$$\min_{V \subseteq \{-1,\cdots,v+2\}, V \ni tx', V \not\ni rx', V \ni rx} \operatorname{rank} H^{lin}_{V,V^c} \geq \dim X_{ax} + d. \tag{3.9}$$

Furthermore, by choosing $V = \{tx', tx, 1, \cdots, v, rx\}$, we have

$$\operatorname{rank} H^{lin}_{\{tx',tx,1,\cdots,v,rx\},rx'} = \operatorname{rank} \begin{bmatrix} A & B_{tx} & B_1 & \cdots & B_v & B_{rx} \end{bmatrix} = \dim X_{ax} + d. \quad (3.10)$$

Therefore, by (3.9) and (3.10) we can conclude

$$\min_{V\subseteq\{-1,\cdots,v+2\}, V\ni tx', V\not\ni rx', V\ni rx} \operatorname{rank} H^{lin}_{V,V^c} = \dim X_{ax} + d.$$

(iii) When $tx \in V$ and $rx \in V^c$,

In this case, we will find a one-to-one mapping between the cutset V for $\mathcal{N}^{lin}(z)$ and a cutset W for $\mathcal{N}(z)$, and show that their mincut is the same with an offset of d.

Let $W := V \setminus \{tx'\}$ and $W' := V^c \setminus \{rx'\} = \{tx', tx, 1, \cdots, v, rx, rx'\} \setminus V \setminus \{rx'\}$. Now, we will show

$$\operatorname{rank} H^{lin}_{V,V^c} = \operatorname{rank} H_{W,W'} + d. \quad (3.11)$$

However, since the proof of (3.11) is not difficult but would be notationally complicated if written out fully, we replace the proof by a representative example. Let $v = 3$ and and $V = \{0, 1\}$.

$$\operatorname{rank} H^{lin}_{V,V^c} = \operatorname{rank} \begin{bmatrix} A & B_0 & B_1 \\ C_4 & 0 & 0 \\ C_2 & 0 & 0 \\ C_3 & 0 & 0 \end{bmatrix}$$

$$\overset{(A)}{=} \operatorname{rank} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I_{d_{rx}} & 0 & 0 & 0 & H_{tx,rx} & H_{1,rx} \\ 0 & 0 & I_{d_{1,out}} & 0 & 0 & H_{tx,1} & H_{1,1} \\ 0 & 0 & 0 & I_{d_{2,out}} & 0 & H_{tx,2} & H_{1,2} \\ 0 & 0 & 0 & 0 & I_{d_{3,out}} & H_{tx,3} & H_{1,3} \\ 0 & -I_{d_{rx}} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -I_{d_{2,out}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -I_{d_{3,out}} & 0 & 0 \end{bmatrix}$$

$$\overset{(B)}{=} \operatorname{rank} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & H_{tx,rx} & H_{1,rx} \\ 0 & 0 & I_{d_{1,out}} & 0 & 0 & H_{tx,1} & H_{1,1} \\ 0 & 0 & 0 & 0 & 0 & H_{tx,2} & H_{1,2} \\ 0 & 0 & 0 & 0 & 0 & H_{tx,3} & H_{1,3} \\ 0 & -I_{d_{rx}} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -I_{d_{2,out}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -I_{d_{3,out}} & 0 & 0 \end{bmatrix}$$

$$\overset{(C)}{=} \operatorname{rank} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & H_{tx,rx} & H_{1,rx} \\ 0 & 0 & I_{d_{1,out}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & H_{tx,2} & H_{1,2} \\ 0 & 0 & 0 & 0 & 0 & H_{tx,3} & H_{1,3} \\ 0 & -I_{d_{rx}} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -I_{d_{2,out}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -I_{d_{3,out}} & 0 & 0 \end{bmatrix}$$

$$\overset{(D)}{=} \operatorname{rank} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & H_{tx,rx} & H_{1,rx} \\ 0 & 0 & 0 & 0 & 0 & H_{tx,2} & H_{1,2} \\ 0 & 0 & 0 & 0 & 0 & H_{tx,3} & H_{1,3} \\ 0 & -I_{d_{rx}} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I_{d_{1,out}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -I_{d_{2,out}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -I_{d_{3,out}} & 0 & 0 \end{bmatrix}$$

$$\overset{(E)}{=} \begin{bmatrix} H_{tx,rx} & H_{1,rx} \\ H_{tx,2} & H_{1,2} \\ H_{tx,3} & H_{1,3} \end{bmatrix} + d$$

$$= \operatorname{rank} H_{W,W'} + d$$

(A): By the definitions of A, B_i, C_i shown in (3.4).
(B): This comes from elementary row operations to eliminate the I's in the A by using the rows in C_i's. In general, this kind of step will make the A part only have I's at the location corresponding to the set V.
(C): This comes from elementary column operations to eliminate the B_i's by using the I's in the A. In general, this kind of step will make the B part to have 0's at the location corresponding to the set V.
(D): By reordering of the rows so that the I's in the A can be grouped with the C_i's. In general, this kind of step will make the B part to be full-rank.
(E): Since we know $\operatorname{rank}\begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix} = \operatorname{rank} A + \operatorname{rank} B$ and by the definitions, $d = d_{rx} + d_{1,out} + d_{2,out} + d_{3,out}$ for this example.

As we can see, we only used elementary row and column operations which hold for general matrices. Thus, we can easily prove that (3.11) holds in general by exactly the above argument.

Finally, using (i),(ii) and (iii) we can prove the lemma.

$$\begin{aligned}
&\min_{V \subseteq \{-1,\cdots,v+2\}, V \ni tx', V \not\ni rx'} \operatorname{rank} H^{lin}_{V,V^c} = \min\{\dim X_{ax}, \min_{W \subseteq \{0,\cdots,v+1\}, W \ni tx, W \not\ni rx} \operatorname{rank} H_{W,W^c}\} + d \\
&= \min\{\dim U, \dim Y, \dim X_{ax}, \min_{W \subseteq \{0,\cdots,v+1\}, W \ni tx, W \not\ni rx} \operatorname{rank} H_{W,W^c}\} + d \\
&= \min\{\operatorname{rank} K_{tx}, \operatorname{rank} K_{rx}, \min_{W \subseteq \{0,\cdots,v+1\}, W \ni tx, W \not\ni rx} \operatorname{rank} H_{W,W^c}\} + d
\end{aligned}$$

Here, the second equality follows from the fact that the mincut of $\mathcal{N}(z)$ is not greater than $\min\{\dim U, \dim Y\}$. The third equality follows from $\operatorname{rank} K_{tx} = \min\{\dim U, \dim X_{ax}\}$ and $\operatorname{rank} K_{rx} = \min\{\dim Y, \dim X_{ax}\}$.

□

The main advantage of linearized networks is that it is known that the algebraic mincut-maxflow theorem holds for $\mathcal{N}_{lin}(z, K)$ [4, Theorem 4.1]. Here, we present the theorem with a simpler, self-contained and different proof for completeness.[4]

Theorem 3.3 (Algebraic Mincut-Maxflow Theorem for Linearized Network [4])**.** *Given the above notations,*

$$\operatorname{rank} G_{lin}(z, K) = \min_{V \subseteq \{-1,\cdots,v+2\}, V \ni tx', V \not\ni rx'} \operatorname{rank} H^{lin}_{V,V^c}$$

Proof. We saw that the transfer functions and channel matrices of $\mathcal{N}_{lin}(z)$ are given in terms of A, B_i, C_i in (3.4) and (3.5) respectively. Thus, it is enough to prove that

$$\operatorname{rank}(A + \sum_{0 \le i \le v+1} B_i K_i C_i) = \min_{V \subseteq \{0,\cdots,v+1\}} \operatorname{rank} \begin{bmatrix} A & B_V \\ C_{V^c} & 0 \end{bmatrix}. \tag{3.12}$$

[4]The proof of [4, Theorem 4.1] only uses linear algebraic fact and relates the rank of the matrices with the rank of bigger matrices. However, here by the use of induction we make each step easier to understand.

This is a fact of linear algebra and can be proved in three steps. First, we prove the theorem for networks with a single relay with a scalar input and output, i.e. $v = -1$ and B_0, C_0 are vectors (Case (i)). Then, we extend the claim for a single relay with a vector input and output, i.e. $v = -1$ and B_0, C_0 are matrices (Case (ii)). Finally, we generalize to multiple relays when $v = 0, 1, 2, \cdots$ (Case (iii)).

(i) First, consider the case when $v = -1$ and B_0, C_0 are vectors i.e. $B_0 \in \mathbb{F}[z]^{m\times 1}$ and $C_0 \in \mathbb{F}[z]^{1\times m}$. Then, (3.12) reduces to

$$\text{rank}(A + B_0K_0C_0) = \min(\text{rank}\begin{bmatrix} A & B_0 \end{bmatrix}, \text{rank}\begin{bmatrix} A \\ C_0 \end{bmatrix}). \tag{3.13}$$

Moreover, since B_0 and C_0 are vectors, $\min(\text{rank}\begin{bmatrix} A & B_0 \end{bmatrix}, \text{rank}\begin{bmatrix} A \\ C_0 \end{bmatrix})$ is either $\text{rank}(A)$ or $\text{rank}(A)+1$.

(i-i) When $\min(\text{rank}\begin{bmatrix} A & B_0 \end{bmatrix}, \text{rank}\begin{bmatrix} A \\ C_0 \end{bmatrix}) = \text{rank}(A)$.

In this case, either $\text{rank}\begin{bmatrix} A & B_0 \end{bmatrix}$ or $\text{rank}\begin{bmatrix} A \\ C \end{bmatrix}$ is equal to $\text{rank}(A)$. Let $\text{rank}\begin{bmatrix} A & B_0 \end{bmatrix} = \text{rank}(A)$. Then, obviously, $\text{rank}(A + B_0K_0C_0) \geq \text{rank}(A)$. Moreover, the column space spanned by B_0 belongs to the column space spanned by A . Thus, $B_0K_0C_0$ cannot increase the rank of the column space and $\text{rank}(A + B_0K_0C_0) = \text{rank}(A)$.

When $\text{rank}\begin{bmatrix} A \\ C \end{bmatrix} = \text{rank}(A)$, the proof follows similarly.

(i-ii) When $\min(\text{rank}\begin{bmatrix} A & B_0 \end{bmatrix}, \text{rank}\begin{bmatrix} A \\ C_0 \end{bmatrix}) = \text{rank}(A) + 1$.

In this case, $\text{rank}\begin{bmatrix} A & B_0 \end{bmatrix} = \text{rank}\begin{bmatrix} A \\ C_0 \end{bmatrix} = \text{rank}(A) + 1$. Moreover, since B_0 is a column vector, $\text{rank}(A + B_0K_0C_0) \leq \text{rank}(A) + 1$. Thus, we only have to prove $\text{rank}(A + B_0K_0C_0) \geq \text{rank}(A) + 1$, which is implied by $\text{rank}(A + B_0C_0) = \text{rank}(A) + 1$. The following claim proves the last statement.

Claim 3.1. *Let $A \in \mathbb{F}[z]^{m\times m}$, $b \in \mathbb{F}[z]^{m\times 1}$, and $c \in \mathbb{F}[z]^{1\times m}$. If*

$$\text{rank}(A) + 1 = \text{rank}\begin{bmatrix} A & b \end{bmatrix} = \text{rank}\begin{bmatrix} A \\ c \end{bmatrix}$$

then

$$\text{rank}(A) + 1 = \text{rank}(A + bc).$$

Proof. Let $\operatorname{rank}(A) = r$. Then, there exist invertible matrices U and V such that

$$UAV = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}.$$

Denote $\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} := Ub$ and $\begin{bmatrix} c_1 & c_2 \end{bmatrix} := cV$ where b_1 and c_1 are $r \times 1$ column and $1 \times r$ row vectors respectively.

Moreover, since U and $\begin{bmatrix} V & 0 \\ 0 & 1 \end{bmatrix}$ are invertible, we have

$$\operatorname{rank} \begin{bmatrix} A & b \end{bmatrix} = \operatorname{rank}(U \begin{bmatrix} A & b \end{bmatrix} \begin{bmatrix} V & 0 \\ 0 & 1 \end{bmatrix}) = \operatorname{rank} \begin{bmatrix} UAV & Ub \end{bmatrix} = \operatorname{rank} \begin{bmatrix} I_r & 0 & b_1 \\ 0 & 0 & b_2 \end{bmatrix} = r + \operatorname{rank}(b_2).$$

Thus, for $\operatorname{rank} \begin{bmatrix} A & b \end{bmatrix} = \operatorname{rank}(A) + 1$ to hold, b_2 has to be a non-zero vector. Likewise, c_2 also has to be a non-zero vector.

Finally, we can conclude

$$\begin{aligned}
\operatorname{rank}(A + bc) = \operatorname{rank}(U(A + bc)V) &= \operatorname{rank}(\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} cV) \\
&= \operatorname{rank}(\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 \\ b_2 \end{bmatrix} cV) && (3.14) \\
&= \operatorname{rank}(\begin{bmatrix} I_r & 0 \\ 0 & b_2 c_2 \end{bmatrix}) && (3.15) \\
&= \operatorname{rank}(A) + 1. && (3.16)
\end{aligned}$$

(3.14): elementary row operation and b_2 is non-zero.
(3.15): elementary row operation.
(3.16): b_2 and c_2 are non-zero. □

(ii) Consider the case when $v = -1$ and B_0, C_0 are general matrices.

Like (i), (3.12) reduces to (3.13). The only difference is now B_0, C_0 can be matrices, and the following claim shows (3.13) still holds.

Claim 3.2. *Let $A \in \mathbb{F}[z]^{m \times m}$, $B_0 \in \mathbb{F}[z]^{m \times r}$, $C_0 \in \mathbb{F}[z]^{q \times m}$, and $K_0 \in \mathbb{F}[K]^{r \times q}$ where each element of K_0 is of the form $k_i \in K$ and distinct. Then,*

$$\operatorname{rank}(A + B_0 K_0 C_0) = \min\{\operatorname{rank} \begin{bmatrix} A & B_0 \end{bmatrix}, \operatorname{rank} \begin{bmatrix} A \\ C_0 \end{bmatrix}\}$$

Proof. Let $x := \text{rank}\begin{bmatrix}A & B_0\end{bmatrix} - \text{rank}(A)$ and $y := \text{rank}\begin{bmatrix}A \\ C_0\end{bmatrix} - \text{rank}(A)$. Then, we can find at least x linearly independent column vectors of B_0 which are independent from the columns of A, and at least y linearly independent row vectors of C_0 which are independent from the rows of A. Formally, let $b_1, \cdots, b_x$ and $c_1, \cdots, c_y$ be such vectors, i.e. b_i and c_j are columns and rows of B_0 and C_0 respectively and $\text{rank}\begin{bmatrix}A & b_1 & \cdots & b_x\end{bmatrix} = \text{rank}\begin{bmatrix}A & B_0\end{bmatrix}$, $\text{rank}\begin{bmatrix}A \\ c_1 \\ \vdots \\ c_y\end{bmatrix} = \text{rank}\begin{bmatrix}A \\ C_0\end{bmatrix}$. Then, we have

$$\text{rank}(A + B_0K_0C_0) \geq \text{rank}(A + \sum_{1\leq i\leq \min\{x,y\}} b_ic_i) \tag{3.17}$$

$$= \min\{\text{rank}\begin{bmatrix}A & B_0\end{bmatrix}, \text{rank}\begin{bmatrix}A \\ C_0\end{bmatrix}\}. \tag{3.18}$$

(3.17): We can find a $r \times q$ matrix K_0' such that all the elements of the matrix are 0 or 1, and $A + B_0K_0'C_0 = A + \sum_{1\leq i\leq\min\{x,y\}} b_ic_i$. Moreover, $\text{rank}(A + B_0K_0C_0) \geq \text{rank}(A + B_0K_0'C_0)$ by Lemma 3.1.

(3.18): b_i and c_i are independent from the column and row space spanned by A respectively. Furthermore, b_i and c_i are also independent from $b_1, \cdots, b_{i-1}$ and $c_1, \cdots, c_{i-1}$ respectively. Therefore, we can repeatedly apply Claim 3.1 and get the desired result.

Moreover,

$$\text{rank}(A + B_0K_0C_0) = \text{rank}(\begin{bmatrix}A & B_0\end{bmatrix}\begin{bmatrix}I \\ K_0C_0\end{bmatrix}) \leq \text{rank}\begin{bmatrix}A & B_0\end{bmatrix} \tag{3.19}$$

$$\text{rank}(A + B_0K_0C_0) = \text{rank}(\begin{bmatrix}I & B_0K_0\end{bmatrix}\begin{bmatrix}A \\ C_0\end{bmatrix}) \leq \text{rank}\begin{bmatrix}A \\ C_0\end{bmatrix}. \tag{3.20}$$

Therefore, by (3.18), (3.19), (3.20) the claim is true. □

(iii) The case with multiple relays, i.e. $v = 0, 1, 2, \cdots$ and B_i, C_i are general matrices.

Now, we will prove (3.12) for a general v. The proof is an induction on $v = -1, 0, 1, 2, \cdots$. Claim 3.2 shows (3.12) is true for $v = -1$. To prove that the theorem also holds for $v = 0, 1, 2, \cdots$, we will assume that the theorem holds for $v = w$ as the induction hypothesis and prove that the theorem holds for $v = w + 1$.

First, by applying Claim 3.2 we have

$$\text{rank}(A + \sum_{0\leq i\leq w+1} B_iK_iC_i) = \text{rank}(A + \sum_{0\leq i\leq w} B_iK_iC_i + B_{w+1}K_{w+1}C_{w+1})$$

$$= \min\{\text{rank}\begin{bmatrix}A + \sum_{0\leq i\leq w} B_iK_iC_i & B_{w+1}\end{bmatrix}, \text{rank}\begin{bmatrix}A + \sum_{0\leq i\leq w} B_iK_iC_i \\ C_{w+1}\end{bmatrix}\} \tag{3.21}$$

Consider the two terms one at a time.

$$\begin{aligned}
&\operatorname{rank}\begin{bmatrix} A + \sum_{0\le i\le w} B_i K_i C_i & B_{w+1}\end{bmatrix} \\
&= \operatorname{rank}(\begin{bmatrix} A & B_{w+1}\end{bmatrix} + \sum_{0\le i\le w} B_i K_i \begin{bmatrix} C_i & 0\end{bmatrix}) \\
&= \min_{W\subseteq\{0,\cdots,w\}} \operatorname{rank}\begin{bmatrix} A & B_{w+1} & B_W \\ C_{W^c} & 0 & 0\end{bmatrix} \qquad (3.22)\\
&= \min_{W\subseteq\{0,\cdots,w+1\}, W\ni w+1} \operatorname{rank}\begin{bmatrix} A & B_W \\ C_{W^c} & 0\end{bmatrix}. \qquad (3.23)
\end{aligned}$$

where (3.22) comes from (3.12) for $v = w$ by replacing A by $\begin{bmatrix} A & B_{v+1}\end{bmatrix}$, B_i by B_i, and C_i by $\begin{bmatrix} C_i & 0\end{bmatrix}$.

Likewise, we can also prove

$$\begin{aligned}
&\operatorname{rank}\begin{bmatrix} A + \sum_{0\le i\le w} B_i K_i C_i \\ C_{w+1}\end{bmatrix} \\
&= \operatorname{rank}\left(\begin{bmatrix} A \\ C_{w+1}\end{bmatrix} + \sum_{0\le i\le w} \begin{bmatrix} B_i \\ 0\end{bmatrix} K_i C_i\right) \\
&= \min_{W\subseteq\{0,\cdots,w+1\}, W\not\ni w+1} \operatorname{rank}\begin{bmatrix} A & B_W \\ C_{W^c} & 0\end{bmatrix} \qquad (3.24)
\end{aligned}$$

By plugging (3.23) and (3.24) to (3.21), we have

$$\operatorname{rank}(A + \sum_{0\le i\le w+1} B_i K_i C_i) = \min_{W\subseteq\{0,\cdots,w+1\}} \operatorname{rank}\begin{bmatrix} A & B_W \\ C_{W^c} & 0\end{bmatrix}.$$

Therefore, by induction the theorem is true. □

So far, we discussed how to convert general topology networks into standardized networks — linearized networks (networks shown in Fig. (3.4) to linearized networks shown in Fig. 3.5). Moreover, we discovered that the mincuts and maxflows of two networks are equivalent with an offset (Lemma 3.3 and Lemma 3.4). Thus, using the mincut-maxflow theorem for linearized networks (Theorem 3.3), we can prove the algebraic mincut-maxflow theorem for general LTI networks.

Proof of Theorem 3.2. Since we can arbitrarily choose d_{ax}, let $d_{ax} \ge max\{d_{tx}, d_{rx}\}$. Then,

$$\begin{aligned}
\operatorname{rank} G(z,K) &= \operatorname{rank}(K_{rx} G(z,K) K_{tx}) && (3.25)\\
&= \operatorname{rank} G_{lin}(z,K) - d && (3.26)\\
&= \min_{V\subseteq\{-1,\cdots,v+2\}, V\ni tx', V\not\ni rx'} \operatorname{rank} H^{lin}_{V,V^c} - d && (3.27)\\
&= \min\{\operatorname{rank} K_{tx}, \operatorname{rank} K_{rx}, \min_{V\subseteq\{0,\cdots,v+1\}, V\ni tx, V\not\ni rx} \operatorname{rank} H_{V,V^c}\} && (3.28)\\
&= \min_{V\subseteq\{0,\cdots,v+1\}, V\ni tx, V\not\ni rx} \operatorname{rank} H_{V,V^c}. && (3.29)
\end{aligned}$$

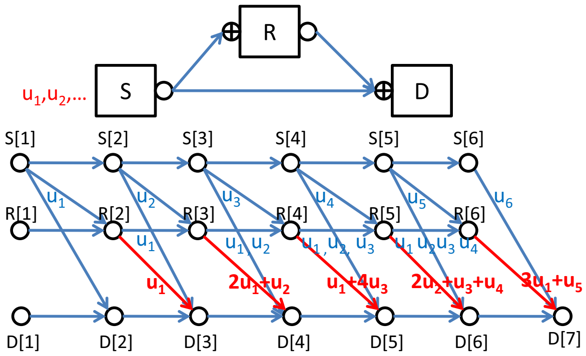


Figure 3.6: Simple Relay Network and the corresponding unfolded network with a mincut-achieving linear scheme

(3.25) is due to the following fact: Select $K_{rx}(z)$ as a $0-1$ matrix that chooses rank $G(z,K)$ independent rows of $G(z,K)$ and $K_{tx}(z)$ as a $0-1$ matrix that chooses rank $G(z,K)$ independent columns of $K_{rx}G(z,K)$. Then, the rank of the resulting matrix $K_{rx}(z)G(z,K)K_{tx}(z)$ is rank $G(z,K)$. Therefore, (3.25) follows from Lemma 3.1.

(3.26), (3.27) and (3.28) follow from Lemma 3.3, Theorem 3.3 and Lemma 3.4 respectively.

(3.29) follows from the fact that the mincut of $\mathcal{N}(z)$ is not greater than $\min\{d_{tx}, d_{rx}\}$, rank $K_{tx} = d_{tx}$ and rank $K_{rx} = d_{rx}$. □

Remark: Part of Theorem 3.2 was already known in [52] and [49]. In fact, the main insight of the theorem is indebted to Koetter and Medard's algebraic framework of network coding [52]. However, the scope of the paper [49] is traditional networks with orthogonal links, and the proof of the theorem is a corollary from the Ford-Fulkerson algorithm [31]. Later, Kim and Medard [49] extended the algebraic framework to the deterministic model [6] using hypergraph ideas, and proved the theorem using Ford-Fulkerson algorithm on hypergraphs [65]. Their idea provides an interesting alternative view to the theorem, and is worth a formal and rigorous study given that the details in [49] were omitted due to space limits. However, the model in [49] is still not general enough for LTI networks since it only covers the case when the channel gains are 0 or 1 and field sizes are finite. Moreover, sometimes it is not clear how to convert general LTI networks to equivalent graphs (or hypergraphs).

3.2.3 Network Linearization vs. Network Unfolding

We proposed network linearization as a way of "converting" an arbitrary relay network to an equivalent acyclic single-hop relay network. In this section, we will compare network linearization with the previously known idea, network unfolding.

Network unfolding is proposed in [1] to convert arbitrary networks to layered networks in which the only existing edges are from one layer to the next layer. As we can see in Fig. 3.6,

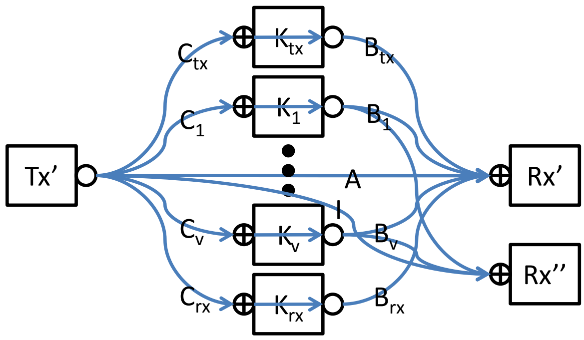


Figure 3.7: Linearized LTI network $\mathcal{N}'_{lin}(Z)$ with an additional destination Rx''

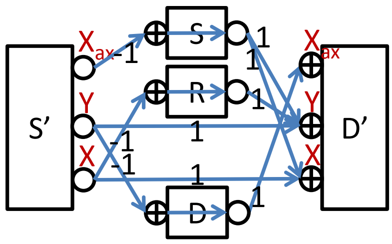


Figure 3.8: Linearized Network of the example in Fig. 3.6

by introducing duplicated nodes over the time, any arbitrary network can be approximated by a layered network. Moreover, the capacity of the layered network approaches the capacity of the original network as the time expansion gets large. Since layered networks have a quite attractive and simple topology, a series of works [36, 2, 112] have exclusively focused on them and developed algorithms that find deterministic linear schemes for layered networks.

However, what these papers are overlooking is that when we fold the unfolded network back into its physical topology a time-invariant scheme might become a time-varying scheme. The example shown in Fig. 3.6 shows that a network-coding design based on an unfolded network can cause significant problems even in the simple network with one source, one relay and one destination. The source transmits $u_1, \cdots, u_6$ to the destination. The letters on the arrows of the unfolded network represent the flows of information. We can easily check that the network-coding scheme shown in the figure is mincut achieving.

However, when we fold it back, we can see problems for implementation. First of all, the scheme is time-varying at the relay. Thus, for the scheme to work every node in the network has to be synchronized to a common clock. Moreover, the transmitted signal at a given time step may depend on all of its previously received signals, which may require a large memory.

On the other hand, from the algebraic mincut-maxflow theorem (Theorem 3.2) we can conclude that there exists a mincut achieving LTI scheme by using the same argument used in [52]. By Lemma 1 of [52] when the field size is large enough there exist K_i that achieve the mincut of the network. Moreover, when the underlying field $\mathbb{F}$ are the reals $\mathbb{R}$ or complex $\mathbb{C}$, these fields already have an infinite number of elements and there exist channel gain matrices which achieve the mincut of the network. When the fields are finite, by extending $\mathbb{F}$ to $\mathbb{F}^m$ we can guarantee a large-enough field size. Furthermore, we even do not have to extend the field when K_{tx}, K_i, K_{rx} are allowed to have memory. $\mathbb{F}[z]$, the field of rational functions in z with coefficients from $\mathbb{F}$, is already an infinite field. Like Lemma 1 of [52] we can prove that there exist mincut-achieving casual[5] LTI filters, K_{tx}, K_i, K_{rx}, whose elements are from $\mathbb{F}[z]$, i.e. having memory is equivalent to extending a field size.

However, we have to be careful to use the network linearization idea for the actual design of the gain matrices K_i, i.e. when we are choosing the elements of K_i from $\mathbb{F}[z]$ and plugging them in. The reason is we also have to guarantee the existence of the transfer function, which is the invertibility of $\begin{bmatrix} I - H_{1,1}K_1 & \cdots & -H_{v,1}K_v \\ \vdots & \ddots & \vdots \\ -H_{1,v}K_1 & \cdots & I - H_{v,v}K_v \end{bmatrix}$ as shown in Theorem 3.1.

Fortunately, this condition can be also posed as a part of the LTI communication network problem. We can easily see that the condition is equivalent to the invertibility of

$$\begin{bmatrix} I_{d_{ax}} & 0 & 0 & \cdots & 0 \\ 0 & I & -H_{1,rx}K_1 & \cdots & -H_{v,rx}K_v \\ 0 & 0 & I - H_{1,1}K_1 & \cdots & -H_{v,1}K_v \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & -H_{1,v}K_1 & \cdots & I - H_{v,v}K_v \end{bmatrix}.$$

This matrix further equals $I + B_1K_1C_1 + \cdots + B_vK_vC_v$ using the definitions in (3.4). We can see the maximum rank (and the dimension) of $I + B_1K_1C_1 + \cdots + B_vK_vC_v$ over all K_i is $d_{ax} + d$. Therefore, the invertibility of the matrix can be thought as the mincut achieving condition from Tx' to Rx'' in Figure 3.7. Finally, we can notice that by choosing d_{ax} as the d.o.f. mincut of $\mathcal{N}(z)$, the maxflow from Tx' to both Rx' and Rx'' becomes $d + d_{ax}$.

Theorem 3.4. *Given the above definitions of $\mathcal{N}(z)$ and $\mathcal{N}'_{lin}(z)$, let's choose d_{ax} as the d.o.f. mincut of $\mathcal{N}(z)$. Then, all the multicast network gains $K_i(z) \in \mathbb{C}^{d_{i,in} \times d_{i,out}}$ which achieve the mincut of $\mathcal{N}'_{lin}(z)$ to both receivers Rx' and Rx'' can also achieve the mincut of $\mathcal{N}(z)$.*

Proof. The proof follows essentially the same as Lemma 3.3 only by replacing K_i with $K_i(z)$. The existence of the transfer function comes from the mincut achievability of Rx'' as discussed above. □

Therefore, we can find a mincut-achieving LTI network coding scheme of $\mathcal{N}(z)$ as follows: (i) Select d_{ax} of (3.3) as the d.o.f. mincut of $\mathcal{N}(z)$. (ii) Find a mincut-achieving multicast network

[5]Notice that even if we put the causal restriction on the design of K_i, the dimensions of the algebraic varieties remain the same. Thus, the proof argument for Lemma 1 of [52] still holds.

coding scheme for the linearized network $\mathcal{N}'_{lin}(z)$ of Figure 3.7 with two receivers. (iii) Apply K_i obtained in the previous steps to the original network.

Furthermore, it is well known that when the network is acyclic, the transfer function always exists [52, Lemma 2]. Therefore, when the network $\mathcal{N}(z)$ is acyclic, the receiver Rx'' in $\mathcal{N}'_{lin}(z)$ which was introduced to guarantee the existence of the transfer function is redundant.

Fig. 3.8 shows the linearized network of the example in Fig. 3.6. By the above argument, any LTI scheme of S, R, D that makes the d.o.f. capacity from S' to D' be 3 achieves the mincut of the original network. For instance, $S = 1$, $R = 1$, $D = 1$ achieve the mincut of both networks of Fig. 3.6 and Fig. 3.8.

Network linearization can also be extended to general information flows, multicast, broadcast, and unicast. Multicast problems will also posed as a multicast problems even after network linearization. However, broadcast and unicast problems will be posed as secrecy problems where eavesdroppers reflect unintended messages in the original problems. We defer further discussions of this to Appendix 8.1.

3.3 Preliminaries on Decentralized Control

In the previous section, we introduced network linearization based on internal states and circulation arcs. As we mentioned, the internal states idea came from linear system theory. Moreover, once we introduce the circulation arc as Fig. 3.4, the whole system becomes a closed-loop system, and such closed-loop systems are the main interest of control theory. Therefore, we can consider control theory from the communication(network coding) perspective. First, we review several known facts on decentralized linear system theory — when the system is LTI-stabilizable — and introduce a few concepts to LTI communication networks.

3.3.1 Decentralized Linear System

Decentralized linear systems have multiple controllers, each of which has access to its own observations and generates its own control inputs. Formally, the decentralized linear system,

$\mathcal{L}(A, B_i, C_i)$, is defined as follows:[6]

$$\begin{aligned} x[n+1] &= Ax[n] + B_1 u_1[n] + \cdots + B_v u_v[n] \\ y_1[n] &= C_1 x[n] \\ &\vdots \\ y_v[n] &= C_v x[n] \end{aligned}$$

where $A \in \mathbb{C}^{m \times m}, B_i \in \mathbb{C}^{m \times q_i}$ and $C_i \in \mathbb{C}^{r_i \times m}$. Then, an interesting question is under what conditions such systems are stabilizable using only LTI controllers:

Definition 3.3 (Stabilizability)**.** *A decentralized linear system is called LTI-stabilizable if there exist linear time-invariant (LTI) controllers $\mathcal{K}_i$ (possibly with internal memories) that connect y_i to u_i whose resulting closed-loop system has only stable poles.*

The stabilizability condition for a decentralized linear system is given in [104] using the concept of fixed modes.

Definition 3.4. *[104, Definition 2] λ is called a fixed mode of $\mathcal{L}(A, B_i, C_i)$ if $\lambda \in \bigcap_{K_i \in \mathbb{C}^{q_i \times r_i}} \sigma(A + \sum_{1 \leq i \leq v} B_i K_i C_i)$ where $\sigma(\cdot)$ is the set of eigenvalues of the matrix.*

The intuition behind this definition is that if an eigenvalue is fixed for all choices of (memoryless) controllers, this eigenvalue is either unobservable or uncontrollable. Thus, if we have unstable fixed modes, we cannot stabilize the plant.

Theorem 3.5. *[104, Theorem 1] $\mathcal{L}(A, B_i, C_i)$ is LTI-stabilizable if and only if all of its fixed modes are within the unit circle.*

Therefore, the stabilizability of linear systems is determined by the existence of unstable fixed modes, and the characterization of stabilizability reduces to characterization of the fixed modes.

However, the characterization of fixed modes shown in Definition 3.4 involves an intersection over an infinite number of sets. Therefore, Anderson *et al.* found the following algebraic characterization of fixed modes (3.1) which only involves minimization over a finite set [4].

Theorem 3.6. *λ is a fixed mode of $\mathcal{L}(A, B_i, C_i)$ if and only if*

$$\min_{V \subseteq \{1,2,\cdots,v\}} \operatorname{rank} \begin{bmatrix} A - \lambda I & B_V \\ C_{V^c} & 0 \end{bmatrix} \geq \dim(A).$$

In other words, the two characterizations of fixed modes shown in Definition 3.4 and Theorem 3.6 are equivalent. In the following discussion, we will see this equivalence turns out to be a special case of the mincut-maxflow theorem for LTI networks.

[6] In this chapter, we consider discrete-time systems since they are conceptually easier to connect to communication theory. We believe that the underlying phenomena discussed here also exist in continuous-time. Furthermore, we assume the matrices here are complex since we will use the Jordan form which can be complex. However, if the system were real we could prove corresponding results restricting the controller design to be real without changing the stabilizability condition.

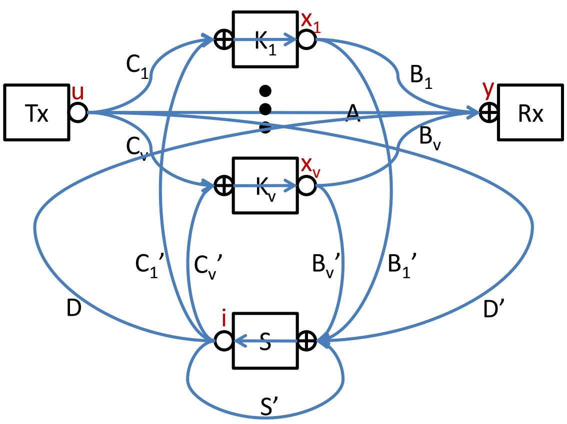


Figure 3.9: Standard LTI Network

3.3.2 LTI Communication Networks at specific frequencies

Since the channel gains of LTI networks are given in z-transform, write the network as $\mathcal{N}(z)$. We will also consider an LTI network, $\mathcal{N}(z)$, at a specific generalized frequency, $z = \lambda$. To indicate that the LTI network is considered at the generalized frequency $z = \lambda$, we write the network as $\mathcal{N}(\lambda)$. $\mathcal{N}(\lambda)$ implies all z in the LTI network are replaced by λ. Then, the capacity definition is naturally generalized to $\mathcal{N}(\lambda)$.

Definition 3.5. *For a given LTI network $\mathcal{N}(z)$, we say that the degree of freedom (d.o.f.) capacity of the network $\mathcal{N}(z)$ is k at frequency $z = \lambda$ if its transfer matrix $G_{tx,rx}(\lambda, K_i)$ is rank k.*

Here we can see that the transfer matrix only makes sense at $z = \lambda$ when it does not have a pole at λ. Thus, we assume that $H_{i,j}$ has no pole at $z = \lambda$. Then, the algebraic mincut-maxflow theorem also holds for $\mathcal{N}(\lambda)$ as before.

Corollary 3.1. *Given the LTI network $\mathcal{N}(\lambda)$ with no poles at λ in the $H_{ij}(z)$,*

$$\begin{aligned}&\mathrm{rank}(G_{tx,rx}(\lambda, K))\\&= \min_{V \subseteq \{tx,1,\cdots,v,rx\}, V \ni tx, V \not\ni rx} \mathrm{rank}(H_{V,V^c}(\lambda)).\end{aligned}$$

Proof. Since the $H_{i,j}(z)$ do not have any pole at λ, we can apply Theorem 3.2 with the channel matrices $H_{i,j}(\lambda)$. □

Before we discuss the externalization of implicit communication in decentralized linear systems, it is helpful to define a standard network we will repeatedly encounter later.

Definition 3.6. *The LTI network shown in Fig. 3.9 is called a standard LTI network,*

$$\mathcal{N}_s(A; B_i, B'_i; C_i, C'_i; D, D'; S, S').$$

The transfer matrix and the channel matrices of the standard network are given as follows.

Lemma 3.5. *In the standard network of Fig. 3.9, the transfer matrix from the transmitter to the receiver is given as*

$$\begin{aligned}G_{tx,rx} =& A + B_1K_1C_1 + \cdots + B_vK_vC_v \\ &+ (D + B_1K_1C_1' + \cdots + B_vK_vC_v') \\ &\cdot (S^{-1} - (S' + B_1'K_1C_1' + \cdots + B_v'K_vC_v'))^{-1} \\ &\cdot (D' + B_1'K_1C_1 + \cdots + B_v'K_vC_v).\end{aligned}$$

The channel matrices H between the transmitter, the relays and the receiver are given for $1 \leq i, j \leq v$:

$$\begin{aligned}H_{tx,rx} &= A + D(S^{-1} - S')^{-1}D', \\ H_{tx,i} &= C_i + C_i'(S^{-1} - S')^{-1}D', \\ H_{i,rx} &= B_i + D(S^{-1} - S')^{-1}B_i', \\ H_{i,j} &= C_j'(S^{-1} - S')^{-1}B_i'.\end{aligned}$$

Here, we just assume the appropriate inverse matrices exist.

Proof. Assign u, x_i, i and y as we can see in Fig. 3.9. Then, we can find the following relationships between these:

$$y = B_1x_1 + \cdots + B_vx_v + Au + Di \tag{3.30}$$

$$x_1 = K_1C_1u + K_1C_1'i \tag{3.31}$$

$$\vdots$$

$$x_v = K_vC_vu + K_vC_v'i$$

$$i = SS'i + SB_1'x_1 + \cdots + SB_v'x_v + SD'u \tag{3.32}$$

By (3.31) and (3.32), we have the following relation:

$$\begin{aligned}&i = SS'i + (SB_1'K_1C_1 + \cdots + SB_v'K_vC_v)u + (SB_1'K_1C_1' + \cdots + SB_v'K_vC_v')i + SD'u \\ &(\Leftrightarrow) S^{-1}i = (S' + B_1'K_1C_1' + \cdots + B_v'K_vC_v')i + (D' + B_1'K_1C_1 + \cdots + B_v'K_vC_v)u \\ &(\Leftrightarrow) i = (S^{-1} - (S + B_1'K_1C_1' + \cdots + B_v'K_vC_v'))^{-1}(D' + B_1'K_1C_1 + \cdots + B_v'K_vC_v)u \quad (3.33)\end{aligned}$$

By plugging (3.31) and (3.33) into (3.30), we get the transfer function from the transmitter to the receiver.

One can easily check the channel matrices between nodes. □

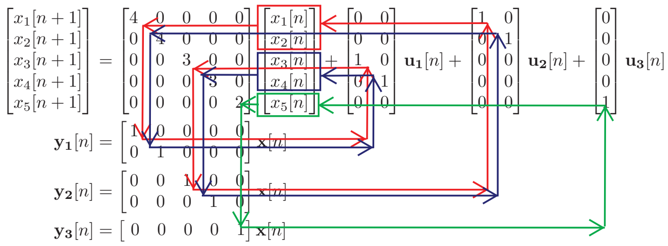


Figure 3.10: An example of an implicit information flow in a decentralized linear system.

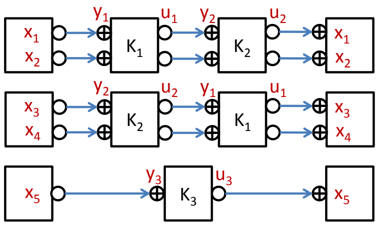


Figure 3.11: Conceptual representations of the information flows within the example of Fig. 3.10

3.4 Example: Information Flow in a Decentralized Linear System

Before we discuss a general algorithm to externalize the implicit communication between controllers, it will be helpful to see the information flows that we want to capture in an illustrative example. By now, we have mounting evidence[7] that in linear systems, the unstable states themselves are the sources and, at the same time, the destinations of information flows. Consider a linear plant controlled by one controller. The states of the system will be excited by the disturbance, *i.e.* the states are generating uncertainties. Then, the states will be observed by the controller, *i.e.* the uncertain information flows from the state to the controller. Finally, the controller will compensate for the disturbance, *i.e.* the information flows back to the states.

When there is more than one controller, the situation becomes more complicated since the controllers can implicitly communicate with each other through the plant [108, 37]. The example shown in Fig. 3.10 (adapted from [5]) illustrates this phenomenon. As we can see, the states $x_1[n]$ and $x_2[n]$ are associated with the eigenvalue 4. However, the controller $\mathcal{K}_1$ can only observe

[7]We return to this point in the conclusion, but the evidence here has largely come from contexts in which the communication is explicitly present. On one side, papers like [89, 26, 86] construct feedback communication systems that use unstable states to encode desired messages. This provides strong evidence for the states acting as information sources. On the other side, papers like [8, 97, 86] talk about networked control systems in which the communication demands on the network come from the states. These argue persuasively for the states in a control system as being destinations of information flows since control and estimation are intimately linked together. The perspective on the Kalman filter presented in [67] suggests strongly that such information flows exist even when there is no explicit communication going on.

$x_1[n], x_2[n]$, the controller $\mathcal{K}_2$ can only control $x_1[n], x_2[n]$, and the controller $\mathcal{K}_3$ can neither observe nor control $x_1[n], x_2[n]$. Therefore, to stabilize $x_1[n], x_2[n]$ the controller $\mathcal{K}_1$ intuitively has to relay its observations to controller $\mathcal{K}_2$ through the implicit channel provided by the states $x_3[n], x_4[n]$.

The red arrow of Fig. 3.10 shows the information flow to stabilize $x_1[n], x_2[n]$. First, $x_1[n], x_2[n]$ is observed by $\mathcal{K}_1$ through $\mathbf{y_1}[n]$. Then, $\mathcal{K}_1$ relays its observations to $\mathcal{K}_2$ by $\mathbf{u_1}[n]$ through the channel $x_3[n], x_4[n]$. $\mathcal{K}_2$ receives the relayed signals through $\mathbf{y_2}[n]$, and finally controls the states by $\mathbf{u_2}[n]$. Thus, we expect that the implicit information flow to stabilize $x_1[n], x_2[n]$ should be roughly representable as the first LTI network of Fig. 3.11. We can see the same kind of information flow to stabilize the states $x_3[n], x_4[n]$ as indicated by the blue arrow. Meanwhile the state $x_5[n]$ can be stabilized by the controller $\mathcal{K}_3$ as indicated by the green arrow. Conceptually, these information flows can be represented as the second and third LTI networks of Fig. 3.11.

Here, we can notice some interesting points. First, we are dividing the states according to their associated eigenvalues. In this example, the states are first divided into three sets $\{x_1[n], x_2[n]\}$, $\{x_3[n], x_4[n]\}$ and $\{x_5[n]\}$, and the information flows for these sets are considered separately. Moreover, in each information flow the states associated with the same eigenvalue are considered as both sources and destinations of the information. The remaining states are considered as the channels that are available to implicitly carry this information flow. The controllers themselves are considered as relays. So in the standard LTI model of Fig. 3.9, the blocks "tx" and "rx" correspond to the set of states in consideration and the remaining states are included in the channel matrices, $A, B_i, \cdots, S'$. The "K_i" blocks correspond to the controllers.

We can also see the connection between stabilizability and capacity. The eigenvalue 4 has two associated states, $x_1[n]$ and $x_2[n]$. Thus, we can think that this source has 2 d.o.f. to transmit. This information can be successfully transferred since the channel provided by the states $x_3[n]$ and $x_4[n]$ has d.o.f. capacity 2, and so the eigenvalue 4 is not a fixed mode. However, if we remove the state $x_4[n]$ from the system, the implicit channel's d.o.f. capacity becomes 1. Thus, a source with 2 d.o.f. cannot be transferred, and the eigenvalue 4 becomes a fixed mode.

Table 3.1 summarizes the relationship between decentralized control and relay communication problems which we have discussed so far and will make rigorous in the following sections.

3.5 Externalization of Implicit Communication

In this section, we discuss how to externalize the implicit communication in decentralized linear systems. The main idea can be considered as the reverse of the algebraic approach to network coding. In [52], Koetter and Medard considered network coding as an algebraic problem. In other words, they found that what is important about networks (graphical objects) in network coding is their transfer functions (algebraic objects). What we do is the opposite. First, we will find transfer functions which are closely connected to the implicit information flows needed to stabilize linear

LTI Communication Networks	Decentralized Linear Systems
Source	Unstable States associated with eigenvalue λ
Destination	Unstable States associated with eigenvalue λ
Relays	Controllers
Channels	Remaining States and B_i, C_i
Message	Unstable Subspace associated with eigenvalue λ
Rate of Message	Number of Jordan blocks associated with eigenvalue λ
Capacity	Stabilizability (Enough implicit communication for unstable subspace)

Table 3.1: The comparison between decentralized linear systems and LTI communication networks

systems. Then, we will find the LTI networks whose transfer functions these are.

3.5.1 Canonical-Form Externalization

It turns out that what is important in externalization is the right choice of transfer function. In section 3.4 we saw that the source and the destination of the information flows are the states. Thus, the straightforward choice is the transfer function from the states $x[n]$ to themselves. For that purpose, we introduce an auxiliary input $u[n]$ and auxiliary output $y[n]$ to the closed loop system in the following way.

$$\begin{aligned} x[n+1] &= (A + B_1K_1C_1 + \cdots + B_vK_vC_v)x[n] + u[n], \\ y[n] &= x[n]. \end{aligned}$$

It is clear that all the states $x[n]$ are directly controllable by $u[n]$ and observable by $y[n]$. Since the fixed modes show up as poles in the transfer function, checking whether λ is a fixed mode involves checking whether the transfer function from $u[n]$ to $y[n]$ has a fixed pole. However, checking poles is mathematically troublesome since it results in division by zero. Thus, instead we inspect the zeros of the formal transfer function from $y[n]$ to $u[n]$.

Under the assumption that $x[0] = 0$, the formal transfer function from $y[n]$ to $u[n]$ is given as

$$u(z) = \underbrace{(zI - A - B_1K_1C_1 - \cdots - B_vK_vC_v)}_{:=G_{cn}(z,K)} y(z).$$

Here, $G_{cn}(z, K)$ is a rational function whose dummy variables are not only z but also the elements of the K_is.

By Lemma 3.5, the standard network, $\mathcal{N}_s(zI - A; -B_i, 0; C_i, 0; 0, 0; 0, 0)$, has $G_{cn}(z, K)$ as its transfer function. Denote this standard network as $\mathcal{N}_{cn}(z)$. The graphical representation of $\mathcal{N}_{cn}(z)$ at the generalized frequency $z = \lambda$ is shown in Fig. 3.12.

Then, we can easily derive the following theorem connecting the d.o.f. capacity of the LTI network $\mathcal{N}_{cn}(z)$ with the stabilizability of the decentralized linear system $\mathcal{L}(A, B_i, C_i)$.

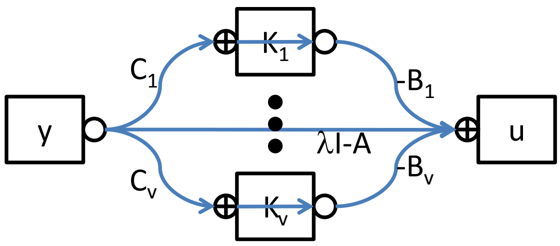


Figure 3.12: The graphical representation of $\mathcal{N}_{cn}(\lambda)$

Theorem 3.7 (Capacity-Stabilizability Equivalence)**.** *Given the above definitions, the following statements are equivalent.*

(1) λ *is a fixed mode of the decentralized linear system* $\mathcal{L}(A, B_i, C_i)$.

(2) $\text{rank}\left(G_{cn}\left(\lambda, K\right)\right) < \dim(A)$.

(3) (transfer matrix rank of LTI network $\mathcal{N}_{cn}(\lambda)$*)* $< \dim(A)$.

(4) (mincut rank of the LTI network $\mathcal{N}_{cn}(\lambda)$*)* $< \dim(A)$.

(5) $\min_{V \subseteq \{1,\cdots,v\}} \text{rank} \begin{bmatrix} \lambda I - A & -B_V \\ C_{V^c} & 0 \end{bmatrix} < \dim(A)$.

Proof. By the definition of fixed modes, (1) is equivalent to $\det(A + \sum_{1 \le i \le v} B_i K_i C_i - \lambda I) = 0$ for all $K_i \in \mathbb{C}^{q_i \times r_i}$. By Lemma 3.1, this is equivalent to $\det(A + \sum_{1 \le i \le v} B_i K_i C_i - \lambda I) = 0$ where each element of K_i is considered as distinct dummy variables. Since $\det(A + \sum_{1 \le i \le v} B_i K_i C_i - \lambda I) = 0$ means not full rank, this is again equivalent to $\text{rank}(\lambda I - A - \sum_{1 \le i \le v} B_i K_i C_i) < \dim(A)$, which is the statement (2). (2) and (3) are equivalent by the definitions of $G_{cn}(z, K)$ and $\mathcal{N}_{cn}(z)$. (3) and (4) are equivalent by the algebraic mincut-maxflow theorem, Corollary 3.1. The equivalence of (4) and (5) follows from the definitions of the cutset matrices of $\mathcal{N}_{cn}(z)$. $\square$

Remark 1: $y(z)$ is the signal assigned to the transmitter of $\mathcal{N}_{cn}(z)$, and $u(z)$ is the signal assigned to the receiver of $\mathcal{N}_{cn}(z)$. Thus, the LTI network connects the states $x[n]$ to themselves, which complies with our discussion of section 3.4.

Remark 2: The statement (1) of the theorem is directly connected to stabilizability by Theorem 3.5, and the statement (3) of the theorem is about the d.o.f. capacity of the network at the frequency $z = \lambda$. Thus, this theorem reveals a fundamental equivalence between stabilizability and capacity.

Remark 3: This externalization seems naive, and as we can see in Fig. 3.12 it gives only networks with a simple topology that does not have any links between the relays. We call this externalization as the canonical-form externalization because of its simple topology. In the next section, we show another way of externalizing the implicit communication which results in a different network topology. The fact that different externalizations are possible is what allowed to us discover that, in fact, any arbitrary network can be converted to the canonical network of Fig. 3.12, which is the insight for network linearization as discussed in Section 3.2.2.

Remark 4: In fact, statement (5) is the algebraic characterization of fixed modes shown in [4]. So

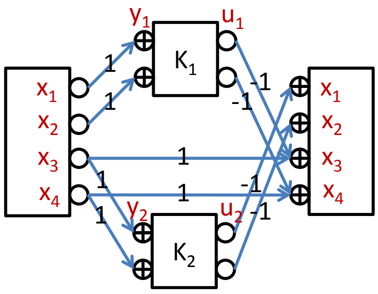


Figure 3.13: Canonical-form externalization of the system of Fig. 3.10 for $\lambda = 4$

in hindsight, we can say that Anderson and Clements found the algebraic mincut-maxflow theorem for the special network of Fig. 3.12.

Remark 5: It is known that the rank of the channel matrix for a cut is a submodular function [112]. The complexity of submodular function minimization is polynomial time [91]. Therefore, we can efficiently check for fixed modes.

Now, we can try to externalize the implicit communication of the example shown in Fig. 3.10. Fig. 3.13 shows the canonical-form externalization for eigenvalue 4. If we look at the figure, this externalization is not what we expected in Fig. 3.11. Since the links between the relays are missing, we cannot see any relaying behavior between two controllers. Also, we cannot clearly see the fact that there are 2 degrees-of-freedom that must be communicated. This motivates us to seek a more compact externalization where the eigenvalues are emphasized by using Jordan forms.

3.5.2 Jordan-Form Externalization

As we see in the above section, externalization is done for each eigenvalue of A. For a general matrix A, there is no clear correspondence between eigenvalues and particular states in the linear system. Thus, we cannot but choose the transfer function from all the states $x[n]$ to themselves. However, if A is given in Jordan normal form [17], we can find a natural correspondence between eigenvalues and states, and use this to reduce the dimension of the transfer function. Moreover, by a similarity transform an arbitrary linear system $\mathcal{L}(A, B_i, C_i)$ can be converted to an equivalent linear system $\mathcal{L}(\widetilde{A}, \widetilde{B_i}, \widetilde{C_i})$ with the matrix A' in Jordan form [17]. Thus, without loss of generality, assume that A is in Jordan form. (This corresponds to examining the system in its natural coordinate system.)

For a Jordan-form A matrix, there is no (internal) interaction between states belonging to different Jordan blocks. Thus, as discussed in section 3.4, to check if λ is a fixed mode, it is enough to examine the transfer matrix from the states associated with Jordan blocks corresponding to the eigenvalue λ to themselves. For externalization, we can simply repeat the steps of the above section.

To understand the core ideas, we first consider a diagonal A matrix, *i.e.* $A = \begin{bmatrix} \lambda I_{m_\lambda} & 0 \\ 0 & A' \end{bmatrix}$ where A' is a diagonal matrix whose diagonal elements are not equal to λ. Because the matrix is diagonal, each Jordan block is just a 1×1 matrix and so m_λ can be thought of as the number of Jordan blocks associated with λ. We will introduce auxiliary inputs and outputs that control and observe the states corresponding to the eigenvalue λ. For this, we define B_λ and C_λ as follows:

$$C_\lambda = \begin{bmatrix} I_{m_\lambda} & 0 \end{bmatrix}, B_\lambda = \begin{bmatrix} I_{m_\lambda} & 0 \end{bmatrix}^T. \tag{3.34}$$

Then, the closed loop system is given as

$$\begin{aligned} x[n+1] &= (A + \sum_{1 \le i \le v} B_i K_i C_i) x[n] + B_\lambda u_\lambda[n] \\ y_\lambda[n] &= C_\lambda x[n] \end{aligned}$$

where $u_\lambda[n]$ and $y_\lambda[n]$ are $m_\lambda \times 1$ vectors. Let's set

$$(zI - A) = \begin{bmatrix} A_{\lambda,1,1}(z) & A_{\lambda,1,2}(z) \\ A_{\lambda,2,1}(z) & A_{\lambda,2,2}(z) \end{bmatrix} \tag{3.35}$$

$$C_i = \begin{bmatrix} C_{i,\lambda,1} & C_{i,\lambda,2} \end{bmatrix}, B_i = \begin{bmatrix} B_{i,\lambda,1} \\ B_{i,\lambda,2} \end{bmatrix}$$

where $A_{\lambda,1,1}(z)$ is a $m_\lambda \times m_\lambda$ matrix, $B_{i,\lambda,1}$ is a $m_\lambda \times q_i$ matrix, $C_{i,\lambda,1}$ is a $r_i \times m_\lambda$ matrix, and the others are the proper implied dimensions. Here, by construction, we can see $A_{\lambda,1,1}(\lambda) = 0$, $A_{\lambda,1,2}(\lambda) = 0$, $A_{\lambda,2,1}(\lambda) = 0$, and $A_{\lambda,2,2}(\lambda)$ is invertible.

Then, we can see that the transfer function from $u_\lambda(z)$ to $y_\lambda(z)$ is given as follows:

$$\begin{aligned} y_\lambda(z) = \begin{bmatrix} I & 0 \end{bmatrix} \Bigg(& \begin{bmatrix} A_{\lambda,1,1}(z) & A_{\lambda,1,2}(z) \\ A_{\lambda,2,1}(z) & A_{\lambda,2,2}(z) \end{bmatrix} \\ & - \sum_{1 \le i \le v} \begin{bmatrix} B_{i,\lambda,1} K_i C_{i,\lambda,1} & B_{i,\lambda,1} K_i C_{i,\lambda,2} \\ B_{i,\lambda,2} K_i C_{i,\lambda,1} & B_{i,\lambda,2} K_i C_{i,\lambda,2} \end{bmatrix} \Bigg)^{-1} \begin{bmatrix} I \\ 0 \end{bmatrix} u_\lambda(z) \end{aligned} \tag{3.36}$$

We need the following lemma to obtain the transfer function from $y_\lambda(z)$ to $u_\lambda(z)$.

Lemma 3.6. *For a field $\mathbb{F}$ and $n_1, n_2 \in \mathbb{Z}^+$, let $y \in \mathbb{F}^{n_1 \times 1}$, $u \in \mathbb{F}^{n_1 \times 1}$, $A \in \mathbb{F}^{n_1 \times n_1}$, $B \in \mathbb{F}^{n_1 \times n_2}$, $C \in \mathbb{F}^{n_2 \times n_1}$, and $D \in \mathbb{F}^{n_2 \times n_2}$. Assume D is invertible. Then, $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ is invertible iff $(A - BD^{-1}C)$ is invertible.*

Moreover, if we assume D and $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ are invertible,

$$y = \begin{bmatrix} I_{n_1} & 0 \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} \begin{bmatrix} I_{n_1} \\ 0 \end{bmatrix} u \tag{3.37}$$

implies

$$u = (A - BD^{-1}C)y.$$

Proof. By Lemma 3.2,

$$\operatorname{rank}\begin{bmatrix} A & B \\ C & D \end{bmatrix} = n_2 + \operatorname{rank}(A - BD^{-1}C).$$

Therefore, the first statement of the lemma is true. For the second,

$$\begin{aligned} y &= \begin{bmatrix} I_{n_1} & 0 \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} \begin{bmatrix} I_{n_1} \\ 0 \end{bmatrix} u \\ &= \begin{bmatrix} I_{n_1} & 0 \end{bmatrix} \left(\begin{bmatrix} I_{n_1} & BD^{-1} \\ 0 & I_{n_2} \end{bmatrix} \begin{bmatrix} A - BD^{-1}C & 0 \\ C & D \end{bmatrix} \right)^{-1} \begin{bmatrix} I_{n_1} \\ 0 \end{bmatrix} u \\ &= \begin{bmatrix} I_{n_1} & 0 \end{bmatrix} \begin{bmatrix} (A - BD^{-1}C)^{-1} & 0 \\ -D^{-1}C(A - BD^{-1}C)^{-1} & D^{-1} \end{bmatrix} \begin{bmatrix} I_{n_1} & -BD^{-1} \\ 0 & I_{n_2} \end{bmatrix} \begin{bmatrix} I_{n_1} \\ 0 \end{bmatrix} u \\ &= (A - BD^{-1}C)^{-1} u \end{aligned}$$

Here, the matrix inverses exist because of the assumption that D is invertible, and the first statement of the lemma. Therefore, $u = (A - BD^{-1}C)y$. □

By Lemma 3.6 and matching (3.36) to the pattern given by (3.37), the transfer function from $y_\lambda(z)$ to $u_\lambda(z)$, $G_{jd,\lambda}(z, K)$, is given as

$$\begin{aligned} G_{jd,\lambda}(z, K) &= (A_{\lambda,1,1}(z) - \sum_{1 \le i \le v} B_{i,\lambda,1} K_i C_{i,\lambda,1}) \\ &+ (A_{\lambda,1,2}(z) - \sum_{1 \le i \le v} B_{i,\lambda,1} K_i C_{i,\lambda,2}) \\ &\cdot (I - (I - A_{\lambda,2,2}(z) + \sum_{1 \le i \le v} B_{i,\lambda,2} K_i C_{i,\lambda,2}))^{-1} \\ &\cdot (-A_{\lambda,2,1}(z) + \sum_{1 \le i \le v} B_{i,\lambda,2} K_i C_{i,\lambda,1}). \end{aligned} \tag{3.38}$$

By Lemma 3.5, $G_{jd,\lambda}(z, K)$ corresponds to the transfer matrix of the standard LTI network,

$$\begin{aligned} &\mathcal{N}_s(A_{\lambda,1,1}(z); -B_{i,\lambda,1}, B_{i,\lambda,2}; C_{i,\lambda,1}, C_{i,\lambda,2} \\ &; A_{\lambda,1,2}(z), -A_{\lambda,2,1}(z); I, I - A_{\lambda,2,2}(z)). \end{aligned} \tag{3.39}$$

Call this network $\mathcal{N}_{jd,\lambda}(z)$. When it is evaluated at the generalized frequency $z = \lambda$, $\mathcal{N}_{jd,\lambda}(z)$ can be simplified further as $\mathcal{N}_s(0; -B_{i,\lambda,1}, B_{i,\lambda,2}; C_{i,\lambda,1}, C_{i,\lambda,2}; 0, 0; I, I - A_{\lambda,2,2}(\lambda))$. Fig. 3.14 shows this

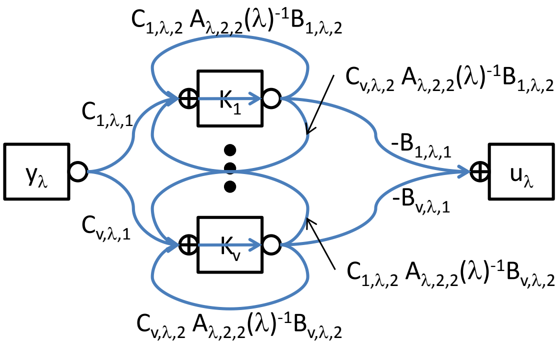


Figure 3.14: The graphical representation of $\mathcal{N}_{jd,\lambda}(\lambda)$

network, $\mathcal{N}_{jd,\lambda}(\lambda)$, and by Lemma 3.5 the channel matrices are given as follows:

$$\begin{aligned}
&H_{tx,rx}(\lambda) = 0, \\
&H_{tx,i}(\lambda) = C_{i,\lambda,1}, \\
&H_{i,rx}(\lambda) = -B_{i,\lambda,1}, \\
&H_{i,j}(\lambda) = C_{j,\lambda,2}A_{\lambda,2,2}(\lambda)^{-1}B_{i,\lambda,2}.
\end{aligned} \tag{3.40}$$

Now, we state a parallel proposition to Theorem 3.7.

Proposition 3.1. *Given the above definitions, the following statements are equivalent.*
(1) λ is a fixed mode of the decentralized linear system $\mathcal{L}(A, B_i, C_i)$
(2) $\text{rank}(G_{jd,\lambda}(\lambda, K)) < m_\lambda$
(3) (transfer matrix rank of LTI network $\mathcal{N}_{jd,\lambda}(\lambda)$) $< m_\lambda$
(4) (mincut rank of the LTI network $\mathcal{N}_{jd,\lambda}(\lambda)$) $< m_\lambda$
(5) $\min_{V\subseteq\{1,\cdots,v\}} \text{rank} \begin{bmatrix} 0 & -B_{V,\lambda,1} \\ C_{V^c,\lambda,1} & C_{V^c,\lambda,2}A_{\lambda,2,2}(\lambda)^{-1}B_{V,\lambda,2} \end{bmatrix} < m_\lambda$

Proof. By Theorem 3.7 (2) and the fact that the dimension of $G_{cn}(\lambda, K)$ is $\dim(A)$, we know that the statement (1) is equivalent to $G_{cn}(\lambda, K)$ is rank deficient. Furthermore, in Lemma 3.6 by considering $(A_{\lambda,1,1}(\lambda) - \sum_{1\le i\le v} B_{i,\lambda,1}K_iC_{i,\lambda,1})$ as A, $(A_{\lambda,1,2}(\lambda) - \sum_{1\le i\le v} B_{i,\lambda,1}K_iC_{i,\lambda,2})$ as B, $(A_{\lambda,2,1}(\lambda) - \sum_{1\le i\le v} B_{i,\lambda,2}K_iC_{i,\lambda,1})$ as C, and $(A_{\lambda,2,2}(\lambda) - \sum_{1\le i\le v} B_{i,\lambda,2}K_iC_{i,\lambda,2})$ as D, we can conclude that $G_{jd,\lambda}(\lambda, K)$ is full rank if and only if

$$\begin{aligned}
&\begin{bmatrix} A_{\lambda,1,1}(z) & A_{\lambda,1,2}(z) \\ A_{\lambda,2,1}(z) & A_{\lambda,2,2}(z) \end{bmatrix} - \sum_{1\le i\le v} \begin{bmatrix} B_{i,\lambda,1}K_iC_{i,\lambda,1} & B_{i,\lambda,1}K_iC_{i,\lambda,2} \\ B_{i,\lambda,2}K_iC_{i,\lambda,1} & B_{i,\lambda,2}K_iC_{i,\lambda,2} \end{bmatrix} \\
&= \lambda I - A - \sum_{1\le i\le v} B_iK_iC_i \\
&= G_{cn}(\lambda, K)
\end{aligned}$$

is full rank. Thus, $G_{cn}(\lambda, K)$ is rank deficient if and only if $G_{jd,\lambda}(\lambda, K)$ is rank deficient. Since the dimension of $G_{jd,\lambda}(\lambda, K)$ is m_λ, the statement (1) is equivalent to the statement (2).

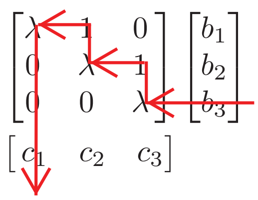


Figure 3.15: Critical Information Flow and Transfer Function in Jordan block

The statement (2) and (3) are equivalent, since $G_{jd,\lambda}(\lambda, K)$ is the transfer function of $\mathcal{N}_{jd,\lambda}(\lambda)$.

The statement (3) and (4) are equivalent by the mincut-maxflow theorem of Corollary 3.1.

The equivalence of the statement (4) and (5) comes from the definitions of the channel matrices of $\mathcal{N}_{jd,\lambda}(\lambda)$ shown in (3.40). □

This theorem can be generalized to arbitrary Jordan forms A by introducing auxiliary inputs and outputs from the states associated with λ to themselves. However, we can further reduce the dimension of the transfer matrix by inspecting the information flow inside nontrivial Jordan blocks.

Let's consider the stabilizability condition for a single Jordan block A matrix, $A = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}$, $B = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$, $C = \begin{bmatrix} c_1 & c_2 & c_3 \end{bmatrix}$. It is well-known [17] that the observability condition for this example is $c_1 \neq 0$ and the controllability condition is $b_3 \neq 0$. In other words, as shown in Fig. 3.15, we can think of the critical information flow to stabilize a single Jordan block as flowing from the right-bottom element to the left-top element. To check whether a single Jordan block has a fixed mode or not, it is enough to consider the transfer function corresponding to this information flow.

This observation for a single Jordan block can be generalized to multiple Jordan blocks. To decide whether λ is a fixed mode or not, it is enough to examine the transfer function matrix from the right-bottom elements of the multiple Jordan blocks (corresponding to the eigenvalue λ) to their left-top elements.

We will make this observation rigorous by introducing the following definitions. Since the definitions are notationally heavy, we recommend visiting Appendix 8.2 for a descriptive example.

In Appendix 8.2, we consider the case when $A = \begin{bmatrix} \lambda & 1 & 0 & 0 & 0 & 0 \\ 0 & \lambda & 1 & 0 & 0 & 0 \\ 0 & 0 & \lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda & 1 & 0 \\ 0 & 0 & 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda' \end{bmatrix}$. Then, we can see that the 3rd and 5th rows and the 1st and 4th column in $\lambda I - A$ are all zeros. To reduce the system to the system considered in Proposition 3.1, we move these all zero columns and rows to left top side of the matrix by multiplying permutation matrices to $\lambda I - A$. To this end, we will define the permutation matrices $P_{L,\lambda}$, $P_{R,\lambda}$.

Let $a_{i,j}$ be the (i, j) element of $A \in \mathbb{C}^{m \times m}$. Since the locations of all zero columns and rows are related to the locations of Jordan blocks, we have to define the indices which indicate the location of each Jordan block. The sequences $\kappa_{L,\lambda}$ and $\kappa_{R,\lambda}$ count the number of Jordan blocks associated with λ. The difference between the two sequences is that $\kappa_{L,\lambda}$ increases at the right-bottom element of the Jordan block, while $\kappa_{R,\lambda}$ increases at the left-top.

$$\kappa_{L,\lambda}(0) = 0$$

For $1 \leq i < m$,

$$\kappa_{L,\lambda}(i) = \begin{cases} \kappa_{L,\lambda}(i-1) + 1 & \text{if } a_{i,i} = \lambda \text{ and } a_{i,i+1} = 0 \\ \kappa_{L,\lambda}(i-1) & \text{otherwise} \end{cases}$$

$$\kappa_{L,\lambda}(m) = \begin{cases} \kappa_{L,\lambda}(m-1) + 1 & \text{if } a_{m,m} = \lambda \\ \kappa_{L,\lambda}(m-1) & \text{otherwise} \end{cases}$$

$$\kappa_{R,\lambda}(0) = 0$$

$$\kappa_{R,\lambda}(1) = \begin{cases} \kappa_{R,\lambda}(0) + 1 & \text{if } a_{1,1} = \lambda \\ \kappa_{R,\lambda}(0) & \text{otherwise} \end{cases}$$

For $1 < i \leq m$,

$$\kappa_{R,\lambda}(i) = \begin{cases} \kappa_{R,\lambda}(i-1) + 1 & \text{if } a_{i,i} = \lambda \text{ and } a_{i-1,i} = 0 \\ \kappa_{R,\lambda}(i-1) & \text{otherwise} \end{cases}$$

Notice that these two sequences are just different ways of counting the number of Jordan blocks associated with the eigenvalue λ. If we denote by m_λ the number of Jordan blocks associated with the eigenvalue λ, then $m_\lambda = \kappa_{L,\lambda}(m) = \kappa_{L,\lambda}(m)$. From the sequences $\kappa_{R,\lambda}$ and $\kappa_{L,\lambda}$, we also define $\iota_{R,\lambda}$ that indicates the left-top elements of the Jordan block associated with λ and $\iota_{L,\lambda}$ that indicates the right-bottom elements.

$$\iota_{L,\lambda}(0) = 0$$
For $1 \le i \le m_\lambda$
$$\iota_{L,\lambda}(i) = \min\{k \in \mathbb{N} : k > \iota_{L,\lambda}(i-1), \kappa_{L,\lambda}(k) > \kappa_{L,\lambda}(k-1)\}$$

Likewise,

$$\iota_{R,\lambda}(0) = 0$$
For $1 \le i \le m_\lambda$
$$\iota_{R,\lambda}(i) = \min\{k \in \mathbb{N} : k > \iota_{R,\lambda}(i-1), \kappa_{R,\lambda}(k) > \kappa_{R,\lambda}(k-1)\}$$

We also define permutation maps and matrices for $\lambda I - A$. The role of these permutation maps and matrices is to collect all zero rows and columns in $\lambda I - A$. The permutation maps $\pi_{L,\lambda}(i)$ and $\pi_{R,\lambda}(i)$ that map the set $\{1, \cdots, m\}$ to itself are defined as follows:

$$\pi_{L,\lambda}(i) = \begin{cases} \kappa_{L,\lambda}(i) & \text{if } \kappa_{L,\lambda}(i) > \kappa_{L,\lambda}(i-1) \\ i + \kappa_{L,\lambda}(m) - \kappa_{L,\lambda}(i) & \text{otherwise} \end{cases}$$
$$\pi_{R,\lambda}(i) = \begin{cases} \kappa_{R,\lambda}(i) & \text{if } \kappa_{R,\lambda}(i) > \kappa_{R,\lambda}(i-1) \\ i + \kappa_{R,\lambda}(m) - \kappa_{R,\lambda}(i) & \text{otherwise} \end{cases}$$

From the permutation map, we define the permutation matrices.

$$P_{L,\lambda} = \begin{bmatrix} e_{\pi_{L,\lambda}(1)} \\ \vdots \\ e_{\pi_{L,\lambda}(m)} \end{bmatrix}, P_{R,\lambda} = \begin{bmatrix} e_{\pi_{R,\lambda}(1)} \\ \vdots \\ e_{\pi_{R,\lambda}(m)} \end{bmatrix} \tag{3.41}$$

where e_i is the row vector with 1 in ith position and 0 in every other position.

Let's multiply these permutation matrices to $zI - A$.

$$P_{L,\lambda}{}^T (zI - A) P_{R,\lambda} = \begin{bmatrix} A_{\lambda,1,1}(z) & A_{\lambda,1,2}(z) \\ A_{\lambda,2,1}(z) & A_{\lambda,2,2}(z) \end{bmatrix} \tag{3.42}$$

where $A_{\lambda,1,1}(z)$ is a $m_\lambda \times m_\lambda$ matrix, $A_{\lambda,1,2}(z)$ is a $m_\lambda \times (m - m_\lambda)$ matrix, $A_{\lambda,2,1}(z)$ is a $(m - m_\lambda) \times m_\lambda$ matrix, $A_{\lambda,2,2}(z)$ is a $(m - m_\lambda) \times (m - m_\lambda)$ matrix.

Since the permutation matrices $P_{L,\lambda}$, $P_{R,\lambda}$ moves all zero columns and rows in $\lambda I - A$ to the left-top side of the matrix (see Appendix 8.2 for an example), we can see $A_{\lambda,1,1}(\lambda) = 0$, $A_{\lambda,1,2}(\lambda) = 0$, $A_{\lambda,2,1}(\lambda) = 0$, and $A_{\lambda,2,2}(\lambda)$ is invertible.

We also multiply the permutation matrices to B_i and C_i, and define the following sub-matrices after this permutation.

$$C_i P_{R,\lambda} = \begin{bmatrix} C_{i,\lambda,1} & C_{i,\lambda,2} \end{bmatrix}, P_{L,\lambda}{}^T B_i = \begin{bmatrix} B_{i,\lambda,1} \\ B_{i,\lambda,2} \end{bmatrix} \tag{3.43}$$

where $B_{i,\lambda,1}$ is a $m_\lambda \times q_i$ matrix, $B_{i,\lambda,2}$ is a $(m - m_\lambda) \times q_i$ matrix, $C_{i,\lambda,1}$ is a $r_i \times m_\lambda$ matrix, $C_{i,\lambda,2}$ is a $r_i \times (m - m_\lambda)$ matrix.

Furthermore, we will also define the auxiliary control and observation matrices B_λ, C_λ as we did in (3.34).

We will introduce an auxiliary input that can control the right-bottom elements of the Jordan blocks and an auxiliary output that can observe the left-top elements of the Jordan blocks. The following matrices B_λ and C_λ correspond to the input and output matrices of the system for these auxiliary input and output.

$$C_\lambda = \begin{bmatrix} e_{\iota_{R,\lambda}(1)} \\ \vdots \\ e_{\iota_{R,\lambda}(m_\lambda)} \end{bmatrix}, B_\lambda = \begin{bmatrix} e_{\iota_{L,\lambda}(1)} \\ \vdots \\ e_{\iota_{L,\lambda}(m_\lambda)} \end{bmatrix}^T.$$

From the construction of the permutation matrices, we can see that when they are applied to C_λ and B_λ, the resulting matrices have nonzero elements only on the left or top side (just as we saw in (3.34)). Formally,

$$C_\lambda P_{R,\lambda} = \begin{bmatrix} I_{m_\lambda \times m_\lambda} & 0 \end{bmatrix}, P_{L,\lambda}^T B_\lambda = \begin{bmatrix} I_{m_\lambda \times m_\lambda} \\ 0 \end{bmatrix}. \tag{3.44}$$

Finally, we get system equations which exactly parallel with the previous diagonal systems in (3.34), (3.35).

Now, we are ready to externalize the implicit communication based on the Jordan form matrix A. Just as the previous diagonal systems, we introduce the auxiliary input $u_\lambda[n] \in \mathbb{C}^{m_\lambda}$ and the auxiliary output $y_\lambda[n] \in \mathbb{C}^{m_\lambda}$. However, unlike the previous section, $u_\lambda[n]$ only controls the right-bottom elements of the Jordan blocks through B_λ and $y_\lambda[n]$ only observes the left-top elements of the Jordan blocks through C_λ.

$$\begin{aligned} x[n+1] &= (A + B_1 K_1 C_1 + \cdots + B_v K_v C_v)x[n] + B_\lambda u_\lambda[n] \\ y_\lambda[n] &= C_\lambda x[n] \end{aligned}$$

Then, the transfer function from $u_\lambda(z)$ to $y_\lambda(z)$ is given as follows:

$$\begin{aligned} y_\lambda(z) &= C_\lambda(zI - A - \sum_{1 \le i \le v} B_i K_i C_i)^{-1} B_\lambda u_\lambda(z) \\ &= C_\lambda \left(P_{L,\lambda} {P_{L,\lambda}}^T \left(zI - A - \sum_{1 \le i \le v} B_i K_i C_i \right) P_{R,\lambda} {P_{R,\lambda}}^T \right)^{-1} B_\lambda u_\lambda(z) \\ &= C_\lambda P_{R,\lambda} \left({P_{L,\lambda}}^T (zI - A) P_{R,\lambda} - \sum_{1 \le i \le v} {P_{L,\lambda}}^T B_i K_i C_i P_{R,\lambda} \right)^{-1} {P_{L,\lambda}}^T B_\lambda u_\lambda(z) \\ &= \begin{bmatrix} I & 0 \end{bmatrix} \left(\begin{bmatrix} A_{\lambda,1,1}(z) & A_{\lambda,1,2}(z) \\ A_{\lambda,2,1}(z) & A_{\lambda,2,2}(z) \end{bmatrix} - \sum_{1 \le i \le v} \begin{bmatrix} B_{i,\lambda,1} \\ B_{i,\lambda,2} \end{bmatrix} K_i \begin{bmatrix} C_{i,\lambda,1} & C_{i,\lambda,2} \end{bmatrix} \right)^{-1} \begin{bmatrix} I \\ 0 \end{bmatrix} u_\lambda(z) \end{aligned} \tag{3.45}$$

where the last line uses (3.44), (3.42), (3.43).

Since (3.36) and (3.45) are the same, (3.38), (3.39), (3.40) still hold. Thus, we can state the capacity-stabilizability equivalence theorem based on the Jordan form A.

Theorem 3.8. *(Capacity-Stabilizability Equivalence 2) Given the above definitions, the following statements are equivalent.*
(1) λ is the fixed mode of the decentralized linear system $\mathcal{L}(A, B_i, C_i)$
(2) $\text{rank}(G_{jd,\lambda}(\lambda, K)) < m_\lambda$
(3) (transfer matrix rank of the LTI network $\mathcal{N}_{jd,\lambda}(\lambda)$) $< m_\lambda$
(4) (mincut rank of the LTI network $\mathcal{N}_{jd,\lambda}(\lambda)$) $< m_\lambda$
(5) $\min_{V \subset \{1,\cdots,v\}} \text{rank} \begin{bmatrix} 0 & -B_{V,\lambda,1} \\ C_{V^c,\lambda,1} & C_{V^c,\lambda,2} A_{\lambda,2,2}(\lambda)^{-1} B_{V,\lambda,2} \end{bmatrix} < m_\lambda$

Proof. The same as Proposition 3.1. □

Remark 1: Notice that the condition (5) seems to be quite different from the statement (5) of Theorem 3.7 that we saw before. However, by remembering that A has Jordan block structure and using the following lemma, we can directly prove the equivalence between these two statements.

Lemma 3.7. *For an invertible square matrix A,*

$$\text{rank} \begin{bmatrix} 0 & 0 & B_0 \\ 0 & A & B_1 \\ C_0 & C_1 & D \end{bmatrix} = \text{rank}\, A + \text{rank} \begin{bmatrix} 0 & B_0 \\ C_0 & D - C_1 A^{-1} B_1 \end{bmatrix}$$

Proof.

$$\text{rank} \begin{bmatrix} 0 & 0 & B_0 \\ 0 & A & B_1 \\ C_0 & C_1 & D \end{bmatrix} = \text{rank} \begin{bmatrix} 0 & B_0 & 0 \\ C_0 & D & C_1 \\ 0 & B_1 & A \end{bmatrix} = \text{rank}\, A + \text{rank} \begin{bmatrix} 0 & B_0 \\ C_0 & D - C_1 A^{-1} B_1 \end{bmatrix}$$

where the first equality is due to elementary row and column operations and the second equality is due to Lemma 3.2. □

Remark 2: This externalization is minimal in the sense that the dimensions of the transmitter input signal and the receiver output signals are minimal. In other words, if we introduce an auxiliary input and output whose dimensions are smaller than the ones shown in this characterization, we cannot find the equivalent condition for fixed modes. The minimality of this characterization manifests as the absence of direct link between the transmitter and the receiver in $\mathcal{N}_{jd,\lambda}(\lambda)$.

Remark 3: It has to be mentioned that this theorem for $m_\lambda = 1$ is already shown in [56]. For this case, the condition (4) of the theorem reduces whether the mincut of the network is 0 or not. Thus, it is equivalent to check the existence of the path from the source to the destination.

The LTI network of Fig. 3.16 shows the Jordan-form externalization of the Fig. 3.10 example for $\lambda = 4$. We can easily see that the LTI network of Fig. 3.16 agrees with the first LTI network

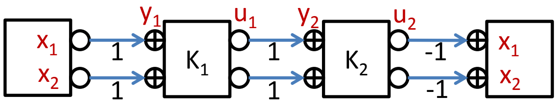


Figure 3.16: Jordan-form externalization of the system of Fig. 3.10 for $\lambda = 4$

of Fig. 3.11. The information generated at $x_1[n], x_2[n]$ is first observed by the controller $\mathcal{K}_1$, then relayed to the controller $\mathcal{K}_2$, and finally returned to $x_1[n], x_2[n]$. Here, the controller $\mathcal{K}_3$ is correctly omitted since it does not affect the transfer function of the relevant LTI network.

Until now, our discussion was limited to strictly proper systems where the impulse response from $u_i[n]$ to $y_j[n]$ is strictly causal. However, the capacity-stabilizability theorem can be easily extended to proper decentralized linear systems $\mathcal{L}(A, B_i, C_i, D_{ij})$ as shown in Appendix 8.3.

Before we close this section, for a sanity check we apply the result of this section to centralized systems which are already well-understood. Moreover, this will be helpful to clarify our understanding in later sections.

Corollary 3.2 (Stabilizability of Centralized Systems[17])**.** *Let's consider the above system with a single controller, $v = 1$. Then, the following conditions are equivalent.*
(1) The centralized linear system $\mathcal{L}(A, B_1, C_1)$ is stabilizable.
(2) (A, B_1) is controllable and (A, C_1) is observable.
(3) $\text{rank}(C_{1,\lambda,1}) \geq m_\lambda$ and[8] *$\text{rank}(B_{1,\lambda,1}) \geq m_\lambda$ for all unstable eigenvalues λ of A.*

Proof. This is a well-known fact in linear system theory [17]. Especially, the equivalence of (1) and (2) immediately follows from Theorem 3.8. □

3.6 Control over LTI networks

To clarify the previous discussion and reveal the further connection between network coding and decentralized linear control, we consider a stabilizability problem with an explicit communication network. Following the problem formulations in [97, 86, 95, 76], we propose 'control over LTI networks' problems. The main advantage of these new problems is that the information for control can only flow explicitly through the communication network, while in general decentralized systems the information can also flow implicitly through the plant. Therefore, we can measure the minimum information flow to stabilize the system by simply measuring the capacity (or reliability) of the explicit communication network.

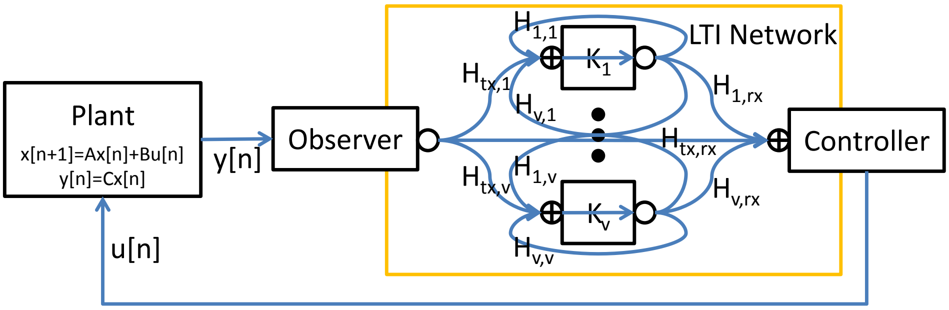


Figure 3.17: Control over LTI Networks: Point-to-Point case

3.6.1 Point-to-Point

The problem of control over LTI networks is shown in Fig. 3.17. The unstable plant is given as

$$x[n+1] = Ax[n] + Bu[n] + w[n]$$
$$y[n] = Cx[n]$$

where $A \in \mathbb{C}^{m \times m}$, $B \in \mathbb{C}^{m \times q_{cn}}$ and $C \in \mathbb{C}^{r_{ob} \times m}$. $x[n]$ is the state, $u[n]$ is the input to the system, $y[n]$ is the output from the system, and $w[n]$ is the disturbance.

The observer can observe the output $y[n]$, but cannot control the plant. On the other hand, the controller can control the plant through the input $u[n]$, but cannot observe the plant. Therefore, to stabilize the plant the observer has to communicate to the controller. The observer and the controller are connected by an LTI communication network, $\mathcal{N}_{ptop}(z)$, where the observer is the transmitter, the controller is the receiver, and the relays are connected by linear time-invariant channels. To make the problem physically meaningful, we assume that the channel matrices $H_{i,j}(z)$ between the relays are stable and causal. Here, we want to find the linear time-invariant observer, controller and relays that stabilize the plant. Therefore, by z-transform, every signal can be represented as a vector in $\mathbb{F}[z]$, and the operation of nodes (controller, observer, and relays) can be represented as a matrix in $\mathbb{F}[z]$. Denote the dimension of the input signal to the LTI network at the observer to be q_{ob}, and that of the output signal from the LTI network at the controller to be r_{cn}. Therefore, the dimensions of the observer and controller gain matrices are $q_{ob} \times r_{ob}$ and $q_{cn} \times r_{cn}$ respectively. At the relay node i, denote the dimension of the input signal to the LTI network to be q_i and that of the output signal from the LTI network to be r_i. Then, the dimension of the relay gain matrix, K_i, is $q_i \times r_i$.

The goal of control and communication nodes is to stabilizing the plant.

Definition 3.7 (Stabilizability over LTI networks)**.** *Given the above definitions, we say the plant is* ***stabilizable over the LTI network*** *if there exist LTI observer, controller and relays that make*

[8] Here, the inequalities are actually equalities, $\text{rank}(C_{1,\lambda,1}) = m_\lambda$ and $\text{rank}(B_{1,\lambda,1}) = m_\lambda$, since the size of $C_{1,\lambda,1}$ and $B_{1,\lambda,1}$ is m_λ.

$x[n]$, $y[n]$, $u[n]$, and all the inputs and outputs of the LTI network uniformly bounded for all uniformly bounded disturbances $w[n]$. For a given design, we say the plant is ***stable over the LTI network*** *if $x[n]$, $y[n]$, $u[n]$, and all the inputs and outputs of the LTI network are uniformly bounded for all uniformly bounded disturbance $w[n]$.*

For a given matrix A, let $\sigma(A)$ be the set of eigenvalues of A. Let m_λ be the number of Jordan blocks of A associated with the eigenvalue λ. Then, the stabilizability condition is given as follows.

Theorem 3.9. *The plant is stabilizable over the LTI network if and only if for all λ such that $\lambda \in \{\lambda : |\lambda| \geq 1\} \cap \sigma(A)$ the following conditions are satisfied:*

$$(i) \begin{bmatrix} \lambda I - A \\ C \end{bmatrix} \text{ is full rank, i.e. } \lambda \text{ is observable.}$$

$$(ii) \begin{bmatrix} \lambda I - A & B \end{bmatrix} \text{ is full rank, i.e. } \lambda \text{ is controllable.}$$

$$(iii) m_\lambda \leq \text{(mincut rank of the LTI network } \mathcal{N}_{ptop}(\lambda))$$

Proof. For the necessity proof, we will use the realization idea. In other words, we will consider control over LTI networks as distributed linear systems and apply the concept of fixed modes to check stabilizability. For the sufficiency proof, we will give a constructive proof. We first design the relays in the LTI network so that it can accommodate enough information flow to stabilize the system. Then, we will design the observer and controller to connect the plant with the communication network, and stabilize it.

(1) Necessity Proof: An insightful reader may notice that 'control over LTI networks' that we are considering is essentially the same as 'decentralized linear systems' of Section 3.3.1. The observer, controller, and relays in Figure 3.17 can be thought of as decentralized controllers. The state $x[n]$ and the internal states of the channels can be combined into one big state $x'[n]$. Then, the minimal realization procedure described in Appendix 8.7 can convert 'control over LTI networks' problems to the following decentralized linear system $\mathcal{L}_{re}(A'_i, B'_i, C'_i, D'_{ij})$.

$$x'[n+1] = A'x'[n] + \sum_{i=0}^{v+1} B'_i u_i[n] + \begin{bmatrix} I_m \\ 0 \end{bmatrix} w[n]$$

$$y_i[n] = C'_i x'[n] + \sum_{j=0}^{v+1} D'_{ij} u_j[n] \text{ for } 0 \leq i \leq v+1$$

Here, the controller 0 and $v+1$ of $\mathcal{L}_{re}(A'_i, B'_i, C'_i, D'_{ij})$ corresponds to the observer and controller of the original problem respectively. The controllers 1 to v correspond to the relays in the original problem. The state $x'[n]$ can be written as $\begin{bmatrix} x[n] \\ x_{ch}[n] \end{bmatrix}$ where $x[n]$ and $x_{ch}[n]$ are respectively the plant and the internal states of the network in the original problem. Then, the state transition matrix A' is a block diagonal matrix $\begin{bmatrix} A & 0 \\ 0 & A_{ch} \end{bmatrix}$.

However, there are minor differences between 'control over LTI networks' and 'decentralized linear system control' problems. In 'control over LTI networks' problems, we only want to stabilize the plant $x[n]$ not all internal states $x'[n]$. And the state disturbance $w[n]$ is also added to only $x[n]$ not to all internal states $x'[n]$. However, since we assume all the channel matrices are stable, the A_{ch} which correspond to $x_{ch}[n]$ have only stable eigenvalues. The only possibly unstable states are $x[n]$. Therefore, by simply repeating the proof shown in [104, 23], we can justify that the stabilizability of the realized system $\mathcal{L}_{re}(A'_i, B'_i, C'_i, D'_{ij})$ is still a necessary condition for stabilizability over the LTI network.

Now, we can apply the Jordan form externalization[9] of Section 3.5.2 for all unstable eigenvalues λ of A. Figure 3.18 shows the resulting LTI network from the Jordan form externalization with respect to λ. By Theorem 3.8, we know that λ is not a fixed mode only if the mincut of the network in Figure 3.18 is greater than m_λ. First, we can think of the cutset that only includes the transmitter y_λ. The channel matrix for this cut is $C_{\lambda,1}$ and so $\operatorname{rank} C_{\lambda,1} \geq m_\lambda$ is a necessary condition for stabilizability. By Corollary 3.2, this is equivalent to the observability of λ which is the condition (i) of the theorem. Likewise, we can think of the cutset that only excludes the receiver u_λ. The channel matrix for this cut is $-B_{\lambda,1}$ and so $\operatorname{rank} B_{\lambda,1} \geq m_\lambda$ is a necessary condition. This corresponds to the theorem's condition (ii), the controllability of λ. The remaining cuts have a one-to-one correspondence to the cuts of the LTI network of Figure 3.17. The conditions that these cuts are larger than m_λ corresponds to the mincut condition of the LTI network, which is the condition (iii) of the theorem.

(2) Sufficiency Proof: For sufficiency, we can also apply the realization idea and use the same sufficiency proof for decentralized linear systems shown in [104, 23]. However, to reveal connections we will give a constructive proof based on network coding, and this style of proof will turn out to be useful in the extensions that we will consider later.

The proof consists of three steps: LTI network design, observer design, and controller design. Without loss of generality, we can assume that A is given in a Jordan form. Then, we can use the notations of Section 3.3.1. For for each unstable eigenvalue λ of A, define the permutation matrices $P_{R,\lambda}$ and $P_{L,\lambda}$ in the same ways as (3.41). Then, we can apply these permutations to the system input and output matrices B and C, and denote the following sub-matrices.

$$C \cdot P_{R,\lambda} = \begin{bmatrix} C_{\lambda,1} & C_{\lambda,2} \end{bmatrix}, P_{L,\lambda}^T \cdot B = \begin{bmatrix} B_{\lambda,1} \\ B_{\lambda,2} \end{bmatrix}$$

where $B_{\lambda,1}$ is a $m_\lambda \times q_{cn}$ matrix, and $C_{\lambda,1}$ is a $r_{ob} \times m_\lambda$ matrix. We will design the controller, observer and relay gain matrices K_{cn}, K_{ob}, K_i. Each element in these gain matrices can be interpreted

[9] Strictly speaking, we have to apply the Jordan form externalization for proper systems shown in Appendix 8.6.

in two ways, ether as a variable in the form of $k_{i,j,k}$, or as constant in $\mathbb{F}[z]$ (i.e. a transfer function given as a z-transform). Then, designing the controller gains can be understood as a procedure of plugging in constants in $\mathbb{F}[z]$ to variables. To distinguish these two meanings of K_i, as mentioned in Section 3.2.1 we will write K_i when it is considered as a variable, and $K_i(z)$ when it is considered as a constant.

(2-a) LTI network (relay) design: The goal of the relays is flowing enough information to stabilize all unstable eigenvalues λ. Denote the transfer function of the LTI network as $G_{ptop}(z, K)$. The goal of the relay gain design is finding $K_i(z) \in \mathbb{F}[z]^{q_i \times r_i}$ such that for all unstable eigenvalues λ, $\text{rank}(G_{ptop}(\lambda, K)) = \text{rank}(G_{ptop}(\lambda, K(z)))$ i.e. achieving the maxflow. Here, because of condition (iii), the maxflow at $z = \lambda$ is always greater or equal to m_λ which is enough to stabilize.

Since the complex (or real) field is infinite, we can find memoryless gains $K_i(z) \in \mathbb{C}^{q_i \times r_i}$ which achieve the maxflow. Rigorously speaking, for each λ, the algebraic variety that makes the rank of $G_{ptop}(\lambda, K)$ be smaller than its maximum rank has a strictly lower dimension than its underlying space. Therefore, there exists an infinite number of solutions that can achieve the maxflow for each λ [52, Lemma 1]. Moreover, even if we have to achieve the maxflow for different eigenvalues simultaneously, the algebraic variety which reduces the rank of any of transfer function matrices just corresponds to a union. Therefore, the dimension is still strictly less than its underlying space, and an infinite number of (simultaneous) solutions exist.

However, when the LTI network has cycles, just guaranteeing the rank condition from the transmitter to the receiver is not enough. Even though all the channel transfer functions are stable, by introducing relay gains at the nodes, we could shift some stable poles to become unstable poles. To prevent such situations, we will adapt the argument introduced by Wang *et al.* in [104]. As shown in [104], using Gershgorin's circle theorem [101] we can prove that as long as the relays gains are chosen small enough, the location of the poles does not move far from the original location. Formally, we can find $\epsilon > 0$ such that for all $|K_i(z)| < \epsilon$ such that $K_i(z) \in \mathbb{C}^{q_i \times r_i}$, all the poles of the LTI network are stable. Moreover, even if we restrict $K_i(z)$ to satisfy $|K_i(z)| < \epsilon$, the dimension of the algebraic variety remains the same. Therefore, the proof of [52, Lemma 1] still holds, and the same argument above guarantees the existence of a mincut achieving $K_i(z)$ which keeps the whole LTI network stable.

(2-b) Observer design: The goal of the observer design is simply connecting all the unstable states of the plant to the LTI network. Mathematically, finding $K_{ob}(z) \in \mathbb{C}^{q_{ob} \times r_{ob}}$ such that for all unstable eigenvalue λ, $\text{rank}(G_{ptop}(\lambda, K(z))K_{ob}C_{\lambda,1}) = \text{rank}(G_{ptop}(\lambda, K(z))K_{ob}(z)C_{\lambda,1})$. Here, we can see since the elements of K_{ob} are variables,

$$\text{rank}(G_{ptop}(\lambda, K(z))K_{ob}C_{\lambda,1}) = \min(\text{rank}(G_{ptop}(\lambda, K(z))), \text{rank}(C_{\lambda,1})).$$

Therefore, by the relay design (2-a) and condition (i) —together with Corollary 3.2— we can con-

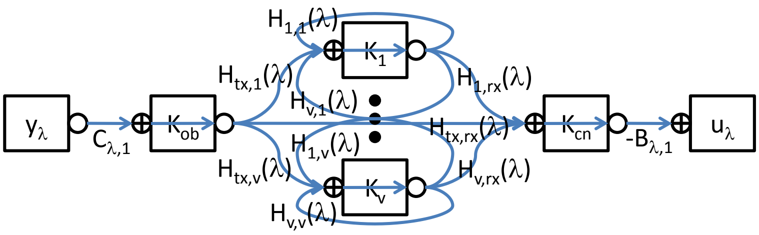


Figure 3.18: Jordan form externalization of $\mathcal{L}_{re}(A'_i, B'_i, C'_i, D'_{ij})$ at $z = \lambda$

clude $\text{rank}(G_{ptop}(\lambda, K(z))K_{ob}C_{\lambda,1}) \geq m_\lambda$. Using the same algebraic variety argument as (2-a), we can prove the existence of such $K_{ob}(z)$. (Here, we do not need Gershgorin's circle theorem for stability.)

(2-c) Controller design: The goal of the controller is to actually stabilize the plant based on the information it got. Once the design of the observer and the relays are fixed, from the controller's point of view the whole system can be viewed as follows in z-transform:

$$zx(z) = Ax(z) + Bu(z)$$
$$y(z) = C'(z)x(z)$$

where $C'(z) = G_{ptop}(z, K(z))K_{ob}(z)C$. For each unstable eigenvalue λ of A, let's apply the same permutation matrix $P_{R,\lambda}$ to $C'(z)$ and denote the following sub-matrices as $C'(z) \cdot P_{R,\lambda} = \begin{bmatrix} C'_{\lambda,1}(z) & C'_{\lambda,2}(z) \end{bmatrix}$. Then, we can easily see $C'_{\lambda,1}(z) = G_{ptop}(z, K(z))K_{ob}(z)C_{\lambda,1}$. Moreover, a simple extension of Corollary 3.2 gives that in this new system, λ is observable if and only if $\text{rank}(C'_{\lambda,1}(\lambda)) \geq m_\lambda$. We already know this condition holds for all unstable eigenvalues λ. Moreover, by condition (ii) all unstable eigenvalues are controllable. Therefore, all unstable eigenvalues are observable and controllable, and so we can stabilize the system using a conventional controller design [17] (which first estimate the states and control the states based on the estimated states).

This finishes the sufficiency proof. □

In the proof of the theorem, we saw how the Jordan form externalization of implicit information flows discussed in Section 3.5.2 can be used to understand problems which have both control and communication aspects. Moreover, the connection between network coding and implicit information flows for control leads to a new controller design for stabilizing the plant.

More importantly, the ideas used in the proof justify our intuition on information flows in decentralized linear system shown in Section 3.4, especially Table 3.1. We converted 'control over LTI networks' problems to decentralized linear systems by considering the relays in LTI networks as controllers of decentralized systems and the channels as a part of the states and input-output matrices B_i, C_i. The goal of the observer and the relays was to send enough information about unstable states associated with λ. Therefore, the unstable states can be considered the source of

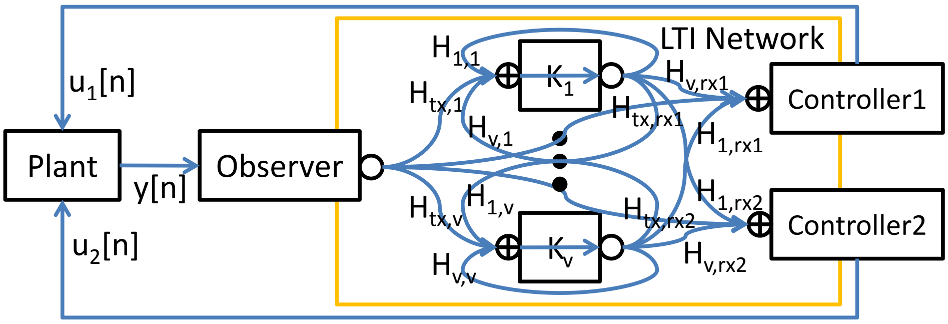


Figure 3.19: Control over LTI Networks with multiple controllers: Multicast case

information flows, and the unstable subspaces can be thought of as the message. The maxflow of the LTI network was compared with m_λ, the number of Jordan blocks associated with λ. Therefore, m_λ can be considered the rate of the message. The controller stabilized the plant by controlling the unstable states based on its received information. Therefore, the unstable states can also be thought of as the destination of information flows. Theorem 3.9 reveals that we can stabilize the system if and only if the LTI network has enough capacity to afford the information flows for control. Therefore, the capacity of LTI networks is deeply related to stabilizability of control systems. Moreover, the communication scheme that we used for the relays was linear network coding.

Another important point is the relationship between network linearization that we discussed in Section 3.2.2 and control over LTI networks. By comparing Figure 3.4 and Figure 3.18, we can easily notice the similarity. The transmitter and receiver in LTI communication networks correspond to the observer and controller in control over LTI networks. These nodes are connected by relay nodes in both problems. Now we can see that what we did by introducing the circulation arc in network linearization (in Figure 3.4) is essentially introducing an unstable plant to be stabilized through the LTI communication network. This insight will be helpful in the later generalization of control over LTI networks, and also the generalization of network linearization in Appendix 8.1.

3.6.2 Multicast

Now, we understand that the distributed controllers communicate by network coding. However, it is known in the communication community that network coding is really helpful to improve the performance when the problem involves multiple transmitters and receivers. Therefore, we will extend the previous single-plant single-observer single-controller problems to problems with multiple plants, observers, and controllers. We will see a close relationship and parallelism between control over LTI networks and network coding.

Arguably, the easiest and most well-understood problem among multi-user network coding problems is the multicast problem. In multicast problems, there is a single transmitter and multiple receivers, and all the receivers want to receive a common message from the transmitter. The worst mincut to all receivers is a trivial lower bound for the message rate in multicast problems. It is shown [1] that we can achieve this lower bound and network coding is necessary for this.

Let's find the counterpart of multicast problems in control over LTI networks. In the sufficiency proof of Thereom 3.9, we saw that the destination of the information flow for control is the controller.[10] Since controllers are the receivers, we have to increase the number of controllers to find the counterpart of multicast problems.

The situation that we will consider in this section is following. Consider a control over LTI networks problem with two controllers as shown in Figure 3.19. Let's say we want to design the system so that the plant becomes stable by either one of the controllers — but does not have to be stable when both controllers are simultaneously active. To design such systems, we can introduce the multicast communication scheme for LTI networks so that the observer sends enough information to stabilize the plant to both controllers.

For simplicity, let's limit our discussion to two controllers but all the results in this section can be easily generalized to multiple controllers. Figure 3.19 shows the resulting problem, control over LTI networks with two controllers. Formally, the plant has two potential control inputs u_1 and u_2, i.e. the plant is given as

$$\begin{aligned} x[n+1] &= Ax[n] + B_1u_1[n] + B_2u_2[n] + w[n] \\ y[n] &= Cx[n] \end{aligned}$$

where $A \in \mathbb{C}^{m \times m}$, $B_1 \in \mathbb{C}^{m \times q_{cn1}}$, $B_2 \in \mathbb{C}^{m \times q_{cn2}}$ and $C \in \mathbb{C}^{r_{ob} \times m}$. If the observations of the observer is decodable at both controllers, it is possible to stabilize the plant by engaging either one. The following definition captures this idea.

Definition 3.8 (Alternative Stabilizability). *Given the above definitions, we say that the plant is* ***alternatively stabilizable over the LTI network*** *if there exist 'common' LTI observer and relays, and possibly different controllers that makes both the first plant*

$$\begin{aligned} x[n+1] &= Ax[n] + B_1u_1[n] + w[n] \\ y[n] &= Cx[n] \end{aligned}$$

and the second plant

$$\begin{aligned} x[n+1] &= Ax[n] + B_2u_2[n] + w[n] \\ y[n] &= Cx[n] \end{aligned}$$

stable over the LTI network.

The reason why this problem is different from just two separate problems with a single controller is that the same observer and relays have to be used for two different systems.

[10] Even if the ultimate destination of the information flow is the unstable states, in control over LTI network problems, only the controller can control the plant. The controller can be thought as a penultimate destination.

Let the LTI network that includes the observer, relays and controller 1 be $\mathcal{N}_{mul1}(z)$. Likewise, the LTI network including the observer, relays and controller 2 is denoted by $\mathcal{N}_{mul2}(z)$. The other notations and assumptions about the problem are the same as the point-to-point case. Then, the condition for alternative stabilizability is given as follows.

Theorem 3.10. *Given the above definitions, the plant is alternatively stabilizable over the LTI network if and only if for all λ such that $\lambda \in \{\lambda : |\lambda| \geq 1\} \cap \sigma(A)$ the following conditions are satisfied*

$$(i) \begin{bmatrix} \lambda I - A \\ C \end{bmatrix} \textit{ is full rank}$$
$$(ii) \begin{bmatrix} \lambda I - A & B_1 \end{bmatrix} \textit{ and } \begin{bmatrix} \lambda I - A & B_2 \end{bmatrix} \textit{ are both full rank}$$
$$(iii) m_\lambda \leq \textit{(mincut rank of the LTI network } \mathcal{N}_{mul1}(\lambda))$$
$$m_\lambda \leq \textit{(mincut rank of the LTI network } \mathcal{N}_{mul2}(\lambda))$$

Proof. (1) Necessity Proof: Since the plant has to be stabilizable by both the controller 1 and 2, the conditions of Theorem 3.9 has to be satisfied for both controllers, which corresponds to the condition (i), (ii), (iii) of the theorem.

(2) Sufficiency Proof: Just as the sufficiency proof of Theorem 3.9, we will give a three-step constructive proof. Since the only difference from that of Theorem 3.9 is LTI network desing, we use the essentially definitions.

(2-a) LTI network design: Since we have to afford enough information flow for both controllers, we choose the relay gain matrices $K_i(z) \in \mathbb{C}^{q_i \times r_i}$ such that for all unstable eigenvalue λ,

$$\text{rank}(G_{mul1}(\lambda, K(z))) \geq m_\lambda \text{ and } \text{rank}(G_{mul2}(\lambda, K(z))) \geq m_\lambda.$$

The existence of such gain matrices can be proved in the same way as Theorem 3.9 and using condition (iii). In other words, the set that we cannot choose $K_i(z)$ is the union of two algebraic varieties: one that makes $G_{mul1}(\lambda, K_i)$ lose its rank and the other one that makes $G_{mul2}(\lambda, K_i)$ lose its rank. The dimension of their union is also strictly smaller than that of the underlying space. Therefore, almost all $K_i(z) \in \mathbb{C}^{q_i \times r_i}$ can achieve the maximum rank of both transfer functions.

(2-b) Observer Design: For the observer design, we find $K_{ob}(z) \in \mathbb{C}^{r_{ob} \times q_{ob}}$ such that for all unstable eigenvalue λ, $\text{rank}(G_{mul1}(\lambda, K(z))K_{ob}(z)C_{\lambda,1}) \geq m_\lambda$ and $\text{rank}(G_{mul2}(\lambda, K(z))K_{ob}(z)C_{\lambda,1}) \geq m_\lambda$. The existence of such $K_{ob}(z)$ follows from the same way as Theorem 3.9 and the union of two algebraic varieties argument.

(2-c) Controller Design: Now, at both controllers the plant is observable. We can simply use conventional controller designs to stabilize the system by both controllers. □

Like the point-to-point problem, memoryless observer and relays are enough for alternative stabilizablity. The generalization of this result to more than two controllers is trivial. We can simply add more controller conditions to the condition (ii) and (iii).

In fact, this theorem can be generalized to arbitrary decentralized linear systems. First, we define strong connectivity of decentralized linear systems [19].

Definition 3.9. *[19] A proper decentralized linear system $\mathcal{L}(A, B_i, C_i)$ with v decentralized controllers is called strongly connected if for all $V \subset \{1, \cdots, v\}$, $C_V(zI - A)^{-1}B_{V^c}$ is nonzero.*

The strong connectivity of the decentralized system implies that for any cut, the transfer function across this cut is not zero. In other word, we can always send some information for any cuts, and thereby every controller is connected with each other.

We generalize the alternative stabilizability definition to a set of decentralized linear systems.

Definition 3.10. *Consider a set of p decentralized linear systems with v decentralized controllers,*

$$\{\mathcal{L}(A^{(1)}, B_i^{(1)}, C_i^{(1)}), \cdots, \mathcal{L}(A^{(p)}, B_i^{(p)}, C_i^{(p)}\}.$$

where for all $2 \leq i \leq v$ the dimensions of $B_i^{(1)}, \cdots, B_i^{(p)}$ are the same, and the dimensions of $C_i^{(1)}, \cdots, C_i^{(p)}$ are also the same.[11] *This set of the decentralized systems is called* ***alternatively stabilizable*** *if there exist common LTI controllers $\mathcal{K}_2, \cdots, \mathcal{K}_v$ and possibly different*[12] *controllers $\mathcal{K}_1^{(1)}, \cdots, \mathcal{K}_1^{(p)}$ such that for all $1 \leq k \leq p$, all systems $\mathcal{L}(A^{(k)}, B_i^{(k)}, C_i^{(k)})$ with controllers $\mathcal{K}_1^{(k)}, \mathcal{K}_2, \cdots, \mathcal{K}_v$ are stable simultaneously.*

The above definition implies that even if the decentralized system is arbitrarily chosen from a given (finite) set, we can stabilize the system by changing only one controller (the controller 1). We can relate this problem with the previous control over LTI network problem. We can consider the observer and relays of control over LTI networks as the controllers 2 through v in decentralized systems. We can consider the multiple controllers as the potential controller 1s in decentralized systems. Therefore, from the realization idea, we can see the alternative stabilization of decentralized linear systems includes that of control over LTI networks as a special case.

This generalized problem corresponds to robust networking [52] in a network coding context. In robust networking, the communication network can be adversarially chosen from a given set, and we want to design the relay scheme that achieves the worst case mincut. In [52], it is shown that robust networking is essentially the same as multicast problems, and the worst case mincut is achievable using network coding.

Likewise, the alternative stabilizability of decentralized linear systems is essentially the same as that of control over LTI networks. If the systems are strongly connected, the alternative stabilizability condition is given as follows.

[11] The dimension of $B_1^{(1)}, \cdots, B_1^{(p)}$ and the dimension of $C_1^{(1)}, \cdots, C_1^{(p)}$ can be different.

[12] the design of the first controller $\mathcal{K}_1^{(i)}$ can be changed depending on which system is chosen.

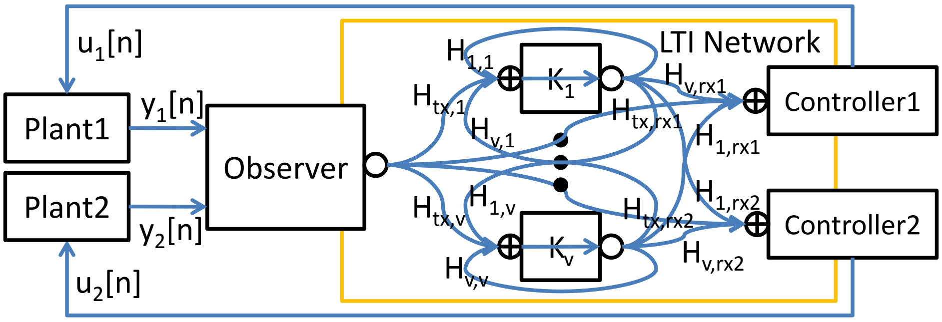


Figure 3.20: Stabilization over an LTI Network with multiple plants and multiple controllers: Broadcast case

Theorem 3.11. *Consider a set of decentralized linear systems with v controllers*

$$\{\mathcal{L}(A^{(1)}, B_i^{(1)}, C_i^{(1)}), \cdots, \mathcal{L}(A^{(p)}, B_i^{(p)}, C_i^{(p)})\}$$

where each decentralized linear system is strongly connected.[13] *Then, this set of decentralized linear systems is alternatively stabilizable if and only if each decentralized linear system does not have unstable fixed modes.*

Proof. The necessity is obvious since each system has to be stabilizable.

Let's prove the sufficiency. By [19, Corollary 1], we know that except a certain algebraic variety whose dimension is strictly smaller than that of the underlying space, almost all constant matrices $K_2(z), \cdots, K_v(z)$ make all unstable eigenvalues of $\mathcal{L}(A^{(1)}, B_i^{(1)}, C_i^{(1)})$ to be observable and controllable at the controller 1. Moreover, by Gershgorin's circle theorem [101], there exists $\epsilon > 0$ such that for all $|K_i(z)| \leq \epsilon$ such that $K_i(z) \in \mathbb{C}^{q_i \times r_i}$, the stable eigenvalues of the system remain stable.

Using the union of algebraic varieties argument, we can prove that there exist constant matrices $K_2(z) \in \mathbb{C}^{q_2 \times r_2}, \cdots, K_v(z) \in \mathbb{C}^{q_v \times r_v}$ such that for all systems $\{\mathcal{L}(A^{(1)}, B_i^{(1)}, C_i^{(1)}), \cdots, \mathcal{L}(A^{(p)}, B_i^{(p)}, C_i^{(p)})\}$, the unstable eigenvalues are observable and controllable at the controller 1 and the stable eigenvalues remain stable. Then, knowing which system is chosen, the first controller can stabilize the system using a conventional design [17]. □

Just as the sufficiency of Theorem 3.10, memoryless controllers are enough for controllers 2 through v. The underlying reason why this theorem holds is that the controllers 2 to v relay (using network coding) enough information for control to controller 1.

3.6.3 Broadcast

Another well-understood problem in network coding is broadcast. Like multicast problems, broadcast problems have a single transmitter and multiple receivers. However, unlike multicast problems, each receiver wants to receiver its own message which is independent from the other's. We can find a simple lower bound on the message rate using cutset bounds. The message rate to receiver 1 cannot exceed the cutset bound for receiver 1, and similar bounds hold for all receivers. We can also think of sum cutsets for augmented receivers. The sum of the message rates to receiver 1 and receiver 2 cannot exceed the cutset bound for the augmented receiver 1 and 2. Likewise, we can think of the cutset bounds for the sum of all two messages, three messages, and so on. This cutset bound is also known to be achievable using network coding together with precoding at the transmitter [62, 52].

In this section, we will find a counterpart of broadcast problems in control over LTI networks. As we saw in the previous section, multiple receivers in network-coding problems correspond to multiple controllers. Now, we have to find the counterpart of multiple messages. In previous discussions, we found that the unstable states correspond to the messages. Therefore, as a counterpart of independent messages, we introduce multiple plants which have orthogonal unstable states. Each controller can only act on its designated plant.

Consider the control over LTI network problem with two plants and two controllers as shown in Figure 3.20. Obviously, we want to design the system so that both plants becomes stable. However, we will require an additional property of disturbance isolation. In other words, if we add disturbance only to plant 1, the states of plant 2 should stay zero for all time. Likewise, if we add disturbance only to plant 2, the states of plant 1 should stay zero for all time. In other words, any disturbance added to the plant 1 must not propagate to plant 2, and vice versa.

For notational simplicity, we will only consider the two plants and two controllers case, but the results in this section can be easily generalized to multiple plants and multiple controllers. Figure 3.20 shows the resulting control over LTI network problem with two plants and two controllers. The plant models are given as follows:

$$x_1[n+1] = A_1 x_1[n] + B_1 u_1[n] + w_1[n]$$
$$y_1[n] = C_1 x_1[n]$$

$$x_2[n+1] = A_2 x_2[n] + B_2 u_2[n] + w_2[n]$$
$$y_2[n] = C_2 x_2[n]$$

where $A_i \in \mathbb{C}^{m_i \times m_i}$, $B_i \in \mathbb{C}^{m_i \times q_{cni}}$, and $C_i \in \mathbb{C}^{r_{obi} \times m_i}$. As shown in Fig. 3.20, the observer has both observations $y_1[n]$ and $y_2[n]$, but both controllers can only control their designated plants via

[13] Otherwise, controllers could be isolated from the remaining system. In this case, for each disconnected system at least one controller's design has to be changed to guarantee stability.

$u_1[n]$ and $u_2[n]$. The basic assumptions and notations for the LTI network are the same as the multicast problem.

If just as in broadcast problems, the observation $y_1[n]$ (information about $x_1[n]$) is decodable separately from $y_2[n]$ at the controller 1 and the observation $y_2[n]$ (information about $x_2[n]$) is decodable separately from $y_1[n]$ at the controller 2, it is possible for controllers to control their own designated plants without causing any interference to the others. This notion is captured by the following definition of independent stabilizablity.

Definition 3.11 (Independent Stabilizability). *Given the above definitions, we say that plants are* ***independently stabilizable over an LTI network*** *if there exist the LTI observer, controllers and relays that satisfy the following conditions:*
(i) both of the plants are stable over the LTI network
(ii) If $w_1[n] = 0$ for all n, then $x_1[n] = 0$ for all n regardless of $w_2[n]$
(iii) If $w_2[n] = 0$ for all n, then $x_2[n] = 0$ for all n regardless of $w_1[n]$

In Figure 3.20, denote the LTI network including the observer, the relays and controller 1 as $\mathcal{N}_{br1}(z)$. Likewise, denote the LTI network that includes the observer, the relays and controller 2 as $\mathcal{N}_{br2}(z)$. The LTI network that has the controller 1 and 2 as the augmented receiver is denoted as $\mathcal{N}_{br1,2}(z)$.

We let $m_{1,\lambda}$ be the number of the Jordan blocks of A_1 associated with the eigenvalue λ, and $m_{2,\lambda}$ be that for A_2. We also let $m_{1,max} := \max_{\lambda \in \mathbb{C}, |\lambda| \geq 1} m_{1,\lambda}$ and $m_{2,max} := \max_{\lambda \in \mathbb{C}, |\lambda| \geq 1} m_{2,\lambda}$.

One may think since we have to prevent disturbance propagation for independent stabilizability, the existence of separate paths from the observer to each controller is required for independent stabilizability. However, we do not need separate paths to each controller. For example, let the plants 1 and 2 be scalar plants. Let the observer have a two-dimensional input signal $\begin{bmatrix} u_{ob,1}[n] \\ u_{ob,2}[n] \end{bmatrix}$ to the network, the controller 1 and 2 have one dimensional $y_{cn1}[n]$ and $y_{cn2}[n]$ respectively, and their relation be given as

$$\begin{bmatrix} y_{cn1}[n] \\ y_{cn2}[n] \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} u_{ob,1}[n] \\ u_{ob,2}[n] \end{bmatrix}.$$

We further assume the network has no relays. In this example, one may think that it is impossible to independently stabilize the system since the communication channels to each controller interfere with each other. However, by simply introducing a precoding gain $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}^{-1}$, we can orthogonalize the paths and independently stabilize the system.

This idea can be formalized for general cases. A sufficient condition and a necessary condition for the independent stabilizability are given as follows.

Theorem 3.12. *Given the above definitions, a sufficient condition for the plants to be independently stabilizable is that for all λ such that $\lambda \in \{\lambda : |\lambda| \geq 1\} \cap (\sigma(A_1) \cup \sigma(A_2))$ the following conditions* *hold:*

$$(i) \begin{bmatrix} \lambda I - A_1 \\ C_1 \end{bmatrix} \text{ and } \begin{bmatrix} \lambda I - A_2 \\ C_2 \end{bmatrix} \text{ are both full rank}$$

$$(ii) \begin{bmatrix} \lambda I - A_1 & B_1 \end{bmatrix} \text{ and } \begin{bmatrix} \lambda I - A_2 & B_2 \end{bmatrix} \text{ are both full rank}$$

$$\begin{aligned} (iii)\ & m_{1,max} + m_{2,max} \leq \text{(mincut rank of the LTI network } \mathcal{N}_{br1,2}(\lambda)) \\ & m_{1,max} \leq \text{(mincut rank of the LTI network } \mathcal{N}_{br1}(\lambda)) \\ & m_{2,max} \leq \text{(mincut rank of the LTI network } \mathcal{N}_{br2}(\lambda)) \end{aligned}$$

The necessary condition for the plants to be independently stabilizable is that for all λ *such that* $\lambda \in \{\lambda : |\lambda| \geq 1\} \cap (\sigma(A_1) \cup \sigma(A_2))$ *the following conditions hold:*

The condition (i) and (ii) hold.

$$\begin{aligned} (iii')\ & m_{1,\lambda} + m_{2,\lambda} \leq \text{(mincut rank of the LTI network } \mathcal{N}_{br1,2}(\lambda)) \\ & m_{1,\lambda} \leq \text{(mincut rank of the LTI network } \mathcal{N}_{br1}(\lambda)) \\ & m_{2,\lambda} \leq \text{(mincut rank of the LTI network } \mathcal{N}_{br2}(\lambda)) \end{aligned}$$

Proof. (1) Necessary condition: The plant 1, the plant 2, and their augmented plant have to be stabilizable by the controller 1, the controller 2, and their augmented controller. Therefore, by apply Theorem 3.9 to these systems, we get the necessary conditions.

(2) Sufficient condition:

The proof is similar to that of Theorem 3.10, but here we need an additional step to remove the interference between the information flows to two controllers. For this, we will adapt the pre-and-post processing idea shown in [62, 52].

(2-a) LTI Network design:

Let $G_{br1}(z, K)$ and $G_{br2}(z, K)$ be the transfer function matrix of $\mathcal{N}_{br1}(z)$ and $\mathcal{N}_{br2}(z)$ respectively. Then, we can see $G_{br1,2}(z, K) := \begin{bmatrix} G_{br1}(z, K) \\ G_{br2}(z, K) \end{bmatrix}$ is the transfer function matrics of $\mathcal{N}_{br1,2}(z)$. Using the same union of algebraic varieties argument of Theorem 3.10, by condition (iii') we can prove that there exist $K_i(z) \in \mathbb{C}^{q_i \times r_i}$ such that for all unstable eigenvalue λ

$$\begin{aligned} & \text{rank}(G_{br1}(\lambda, K(z)) \geq m_{1,max} \\ & \text{rank}(G_{br2}(\lambda, K(z)) \geq m_{2,max} \\ & \text{rank}(G_{br1,2}(\lambda, K(z))) \geq m_{1,max} + m_{2,max} \end{aligned} \tag{3.46}$$

and keep the stable eigenvalues stable.

(2-b) Pre-and-Post processors at Controller and Observer: Even if we design the relays so that they can flow enough information, information flows from the observer to the controllers can interfere with each other. To remove this interference, we introduce pre-and-post processors at the controllers and observer as shown in [62, 52].

First, let's make $G_{br1,2}(z,K(z))$ a square matrix by introducing pre-and-post processors $K'_{cn1}(z) \in \mathbb{C}^{m_{1,max} \times r_{cn1}}$, $K'_{cn2}(z) \in \mathbb{C}^{m_{2,max} \times r_{cn2}}$, $K'_{ob}(z) \in \mathbb{C}^{q_{ob} \times (m_{1,max}+m_{2,max})}$ as follows:

$$G'_{br1,2}(z,K(z)) := \begin{bmatrix} K'_{cn1}(z) & 0 \\ 0 & K'_{cn2}(z) \end{bmatrix} G_{br1,2}(z,K(z)) K'_{ob}(z).$$

The resulting matrix $G'_{br1,2}(z,K(z))$ is a square matrix with dimension $(m_{1,max} + m_{2,max})$, and using the algebraic variety argument and (3.46) we can choose $K'_{cn1}(z)$, $K'_{cn2}(z)$, $K'_{ob}(z)$ so that for all unstable eigenvalues λ, $G'_{br1,2}(\lambda, K(z))$ is invertible.

Now, we can remove the interference by simply multiplying by the matrix inverse. To this end, denote

$$K''_{ob}(z) := z^{-d} det(G'_{br1,2}(z,K(z))) G'_{br1,2}(z,K(z))^{-1}$$

Here, we introduce z^{-d} to make $K''_{ob}(z)$ causal. Therefore, $d \in \mathbb{Z}^+$ has to be chosen large enough so that each element in $K''_{ob}(z)$ is causal. Furthermore, since we multiplied $det(G'_{br1,2}(z,K(z)))$, $K''_{ob}(z)$ does not have any additional pole beyond the existing ones in $G'_{br1,2}(z,K(z))$. Thus, $K''_{ob}(z)$ is also stable. Let's multiply this matrix to $G'_{br1,2}(z,K(z))$ and denote

$$G''_{br1,2}(z,K(z)) := G'_{br1,2}(z,K(z)) K''_{ob}(z).$$

In $G''_{br1,2}(z,K(z))$, the only non-zero entries are diagonal entries, and so we have $(m_{1,max} + m_{2,max})$ "orthogonal" communication channels.

(2-c) Observer design: In the observer, we will use $m_{1,max}$ communication channels to send information about plant 1, and the remaining for plant 2. First, denote $C_{1,\lambda,1}$ and $C_{2,\lambda,1}$ for C_1 and C_2 in the same way we defined $C_{\lambda,1}$ for C in Theorem 3.10. Using the algebraic variety argument from Theorem 3.10 and condition (i), we can show that there exist $K'''_{ob}(z) \in \mathbb{C}^{m_{1,max} \times q_{cn1}}$ and $K''''_{ob}(z) \in \mathbb{C}^{m_{2,max} \times q_{cn2}}$ such that for all unstable eigenvalues λ,

$$\text{rank}(K'''_{ob}(z) C_{1,\lambda,1}) \geq m_{1,\lambda}$$
$$\text{rank}(K''''_{ob}(z) C_{2,\lambda,1}) \geq m_{2,\lambda}.$$

Then, we will set the observer gain $K_{ob}(z)$ as

$$K_{ob}(z) = K'_{ob}(z) K''_{ob}(z) \begin{bmatrix} K'''_{ob}(z) & 0 \\ 0 & K''''_{ob}(z) \end{bmatrix}.$$

(2-d) Controller design: Once we fix the relay gain and observer gain matrices as above and introduce the gain matrix $K'_{cn1}(z)$ at controller 1, by construction the controller 1 will have the following observation about the state.

$$
\begin{aligned}
&\begin{bmatrix} K'_{cn1}(z) & 0 \end{bmatrix} G_{br1,2}(z, K(z)) K'_{ob}(z) K''_{ob}(z) \begin{bmatrix} K'''_{ob}(z) & 0 \\ 0 & K''''_{ob}(z) \end{bmatrix} \begin{bmatrix} y_1(z) \\ y_2(z) \end{bmatrix} \\
&= \begin{bmatrix} K'_{cn1}(z) & 0 \end{bmatrix} G_{br1,2}(z, K(z)) K'_{ob}(z) K''_{ob}(z) \begin{bmatrix} K'''_{ob}(z) & 0 \\ 0 & K''''_{ob}(z) \end{bmatrix} \begin{bmatrix} C_1 x_1(z) \\ C_2 x_2(z) \end{bmatrix} \\
&= z^{-d} \det(G'_{br1,2}(z)) \begin{bmatrix} I & 0 \end{bmatrix} \begin{bmatrix} K'''_{ob}(z) & 0 \\ 0 & K''''_{ob}(z) \end{bmatrix} \begin{bmatrix} C_1 x_1(z) \\ C_2 x_2(z) \end{bmatrix} \\
&= z^{-d} \det(G'_{br1,2}(z)) K'''_{ob}(z) C_1 x_1(z)
\end{aligned}
$$

As we can see, the observation is orthogonal to the state of plant 2. Moreover, since for all unstable eigenvalue λ, $\det(G'_{br1,2}(\lambda)) \neq 0$ and $K'''_{ob}(z)C_1$ can observe all unstable states of $x_1[n]$, the plant 1 is observable. Therefore, by a conventional controller design, controller 1 can orthogonally stabilize plant 1. The same holds for plant 2 and controller 2. □

The result can be easily generalized to multiple plants and multiple observers. Unlike Theorem 3.9 and Theorem 3.10, the memories at the observer and the relies are actually helpful. The necessary and the sufficient condition coincide when all the unstable eigenvalues of A_1 and A_2 are the same, and this corresponds to the broadcast result of network coding.

However, unlike broadcast problems in network coding, the augmentation idea of nodes and cutset bounds fail to give a tight necessary condition. The reason for this is in this problem we have an additional factor, the frequency z. According to the frequency where it is evaluated, the channel behaves significantly differently. Thus, there is no way to orthogonalize the channel simultaneously for all frequencies, and we cannot achieve the necessary condition obtained by the augmentation idea.

For example, let's consider the plant $A_1 = 3$, $A_2 = 2$, $B_1 = B_2 = 1$ and $C_1 = C_2 = 1$. And the LTI network has no relays, the input signal dimension of the observer and the output signal dimension of the controllers are 1, and $G_{br1,2}(z, K_i) = \begin{bmatrix} 3 - 6z^{-1} \\ 2 - 6z^{-1} \end{bmatrix}$. Here, since there are two scalar plants and the observer has only one dimensional input signal to the network, it "seems impossible" to independently stabilize the systems. In fact, this system violates the sufficiency condition of Theorem 3.12 since $m_{1,max} = 1$, $m_{2,max} = 1$, and the mincut ranks of $\mathcal{N}_{br1}(3)$, $\mathcal{N}_{br2}(2)$ are both 1. Therefore, Theorem 3.12 fails to guarantee independent stabilizability of the system.

However, the system still satisfies the necessary condition of Theorem 3.12 derived by a simple augmented system idea. We can easily check that the system parameters are $m_{1,3} = 1$, $m_{2,3} = 0$, (mincut rank of $\mathcal{N}_{br1,2}(3)$)=1, (mincut rank of $\mathcal{N}_{br1,2}(3)$)=1, (mincut rank of $\mathcal{N}_{br1,2}(3)$)=0, $m_{1,2} = 0$, $m_{2,2} = 1$, (mincut rank of $\mathcal{N}_{br1,2}(2)$)=1, (mincut rank of $\mathcal{N}_{br1,2}(2)$)=0, (mincut rank of $\mathcal{N}_{br1,2}(2)$)=1. These parameters satisfy the necessary condition of the theorem.

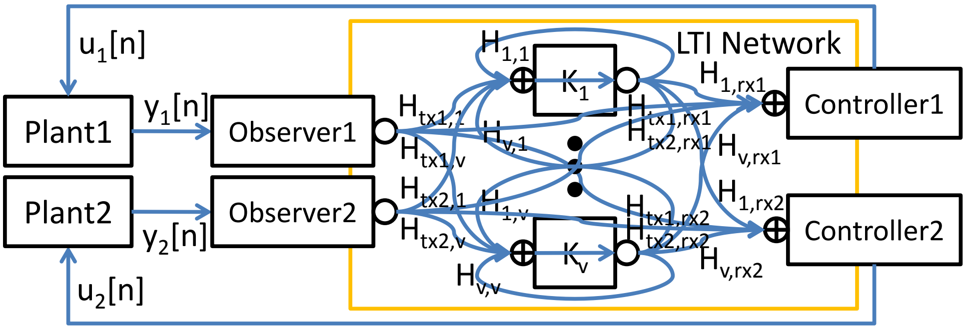


Figure 3.21: Control over LTI Networks with multiple plants, multiple observers, and multiple controllers: Multiple-unicast case

Therefore, for even for this simple system, the necessary and sufficient condition of Theorem 3.12 do not match. Finding the tight characterization for the independent stabilizability will be an interesting further research direction. The example shows that it is the necessity condition that probably needs to be tightened.

3.6.4 Multiple-Unicast

Multiple-unicast problems in network coding have multiple transmitter-receiver pairs which try to communicate their own individual messages. Unlike the previous problems, each transmitter only knows its own messages, and it is well-known that the cutset bound is not tight and the capacity region is open except several known cases [102, 103].

Here, we try to convert multiple-unicast problem to the control over LTI network problems. The main difference between multiple-unicast and broadcast problems is the multiple transmitters. To capture this, we will introduce multiple observers[14] to the previous control over LTI network problems.

Figure 3.21 shows the resulting problem. The only difference compared with Figure 3.19 is the multiple observers which do not share their observations directly. In this problem, we can easily prove that if there exist a multiple unicast communication scheme from the observers to the controllers which accommodates enough information flow to stabilize the plants, we can independently stabilize the system.

3.7 Conclusion

In this chapter, we take a unified approach to network coding and decentralized control by considering both problems as linear time-invariant systems. LTI-stabilizability of decentralized

[14]In section 3.4, we argued that the sources of the information flows for control are unstable states. However, when only explicit observers can directly observe the unstable states, the observers can be thought of as the sources of information.

linear systems is found to be equivalent to having sufficient capacity in the relevant LTI networks. This equivalence can be exploited in both network coding and decentralized control contexts.

In network coding, we found network linearization by introducing internal states and circulation arcs. The linearized network has not only an equivalent mincut and maxflow to the original network, but also a simple topology, acyclic single-hop relay. These properties lead to a simple and elegant proof of an algebraic mincut-maxflow theorem.

In decentralized control, we gave an algorithm to make explicit LTI communication networks that represent the implicit communication required to stabilize the plant. The stabilizability condition of decentralized systems is then easily interpreted using mincut conditions on the corresponding networks. Each eigenvalue is viewed separately, and the number of Jordan blocks corresponding to that eigenvalue corresponds to the number of degrees-of-freedom of implicit communication required to stabilize that eigenvalue. The algebraic condition for fixed modes that was reported in [4] and had, in our opinion, remained mysterious for 30 years turns out to be a special case of the algebraic mincut-maxflow theorem. This also confirms that LTI controllers in decentralized control systems implicitly communicate via linear network coding.

The connection to network coding becomes even more clear when we consider stabilization problems with an explicit communication network. By introducing the concepts of alternative stabilizability and independent stabilizability, we successfully convert network-coding results to equivalent stabilizability results.

Taking a step back, the general idea of implicit communication (signaling) between decentralized controllers and information flow in decentralized systems has been recognized since Witsenhausen's counterexample [108]. However, in Witsenhausen's counterexample the need for communication between controllers is justified by the suboptimality of linear controllers, *i.e.* if the decentralized controllers want to communicate with each other for efficient control of the system, they would do so using nonlinear controllers for signaling [109, 45, 37]. However, we showed here that even if we restrict controllers to be linear time-invariant, the controllers still can communicate via linear network coding. To an extent, this chapter does for implicit communication what [95, 27] did vis-a-vis [86, 87] for explicit communication — it finds a way to discuss the issue within a linear framework. In fact, the existence of implicit communication between linear controllers in decentralized systems has been conjectured for a long time [5, 19, 3, 116]. In a sense, we hope that this chapter clarifies these discussions.

Chapter 4

Decentralized scalar LQG problem: Fast Dynamics

4.1 Introduction

One of the biggest successes in stochastic control theory is the LQG (linear quadratic Gaussian) problem with a single controller. The solution of the LQG problem contributed two big ideas to classical control theory [55]: The first is the optimality of linear controllers. This fact allows designers to confidently focus on finite-dimensional linear strategies without worrying about the infinite-dimensional strategy space. The second is the optimality of the Certainty-Equivalent-Controllers (CEC). Without loss of optimality, we can first estimate states and then control the system as if the estimated states were the true states. This is also called the estimation and control separation principle.

Even if the optimality results were restricted to single-controller LQG problems, their philosophical contribution was not limited to them. Lots of related but different control areas — including nonlinear system control and adaptive control — accepted these principles and focused on essentially linear controllers, and separated estimation from control. In this sense, the LQG problems form a conceptual foundation in control theory.

However, this beautiful result on the LQG problem with a single controller fails as soon as we introduce more than one controller. Following convention, we call a problem with a single controller a centralized problem, and one with multiple controllers a decentralized problem. The famous Witsenhausen's counterexample [108] demonstrates that nonlinear strategies outperform linear strategies even in a simple finite-horizon decentralized LQG problem. Later, Ho, Kastner, and Wong [45] qualitatively argued that the need for nonlinear controllers stems from "signaling" — we will also use the term "implicit communication" interchangeably — between decentralized controllers. Finding the optimal nonlinear strategy in most decentralized problems is known to be a non-convex infinite-dimensional problem [88], for which we do not have a well-developed theory.

Yet, it is still interesting to consider the average-cost infinite-horizon decentralized LQG problem, which is the natural extension of [55, p.93].

$$\mathbf{x}[n+1] = \mathbf{A}\mathbf{x}[n] + \sum_i \mathbf{B_i}\mathbf{u_i}[n] + \mathbf{w}[n]$$

$$\mathbf{y_i}[n] = \mathbf{C_i}\mathbf{x}[n] + \mathbf{v_i}[n]$$

Here, the underlying random variables $\mathbf{x}[0]$, $\mathbf{w}[n]$, and $\mathbf{v_i}[n]$ are independent Gaussian. The objective is to minimize the asymptotic average cost:

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{0\leq n<N} \mathbb{E}[\mathbf{x}^*[n]\mathbf{Q}\mathbf{x}[n]] + \sum_i \mathbb{E}[\mathbf{u_i}^*[n]\mathbf{R_i}\mathbf{u_i}[n]]$$

where $\mathbf{Q} \succeq \mathbf{0}$, $\mathbf{R_i} \succeq \mathbf{0}$, and each $\mathbf{u_i}[n]$ is a causal function of $\mathbf{y_i}[n]$ alone. This chapter (and the next) considers the simplest toy case among these infinite-horizon decentralized LQG problems, a scalar system with two-controllers. As should be expected, linear controllers are not optimal. The crux of decentralized LQG problems, nonconvex optimization over infinite-dimensional spaces, is still there and finding the optimal solution seems impossible. Instead of trying to solve the problem exactly, we solve it **approximately** to within a constant factor of the optimal cost.

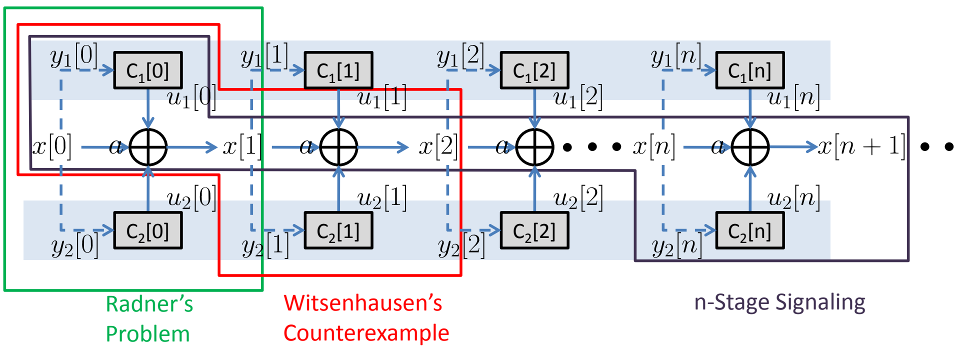


Figure 4.1: Relationship between Radner's problem [80], Witsenhausen's counterexample [108], and the infinite-horizon scalar LQG problem with two controllers

4.1.1 Literature Review and Intellectual Context

Until the late 40s, control and communication were considered in a unified framework under the name of *cybernetics*. According to Wiener [105], cybernetics is defined as 'the scientific study of control and communication in the animal and the machine.' However, Shannon's revolutionary paper detached communication problems as its own field of interest. Since then, control and communication have grown as two separate areas.

Now, control theory which successfully addressed fundamental problems in centralized control is facing the decentralized challenge (non-convex infinite-dimensional optimization problems). This challenge divided related control research in two major directions.

The first direction is finding those special cases under which a linear strategy *is* optimal — or almost equivalently, finding cases where the problem is convex. Radner's pioneering paper [80] considered the case when the controllers act simultaneously and the dynamics of the system terminates after one time step, as shown in the first box of Fig. 4.1. Signaling (implicit communication between two controllers) is intuitively impossible by problem construction. Therefore, linear controllers are optimal in this case in spite of the problem being decentralized. Witsenhausen found another sufficient condition for linear optimality called the nested information pattern [109]. The condition tells if all information is shared with one step time delay by explicit communication, there is no need to implicitly communicate the information and linear strategies are optimal. Later, this concept was generalized by Yuksel to stochastic nestedness [117].

More recently, Rotkowitz and Lall [84] proposed an algebraic condition for convexity of the problem called "quadratic invariance." The condition finds sparsity constraints on the controller so that the problem remains convex even after Youla's parametrization [114]. There is a lot of ongoing research in this direction [92, 60] that has refined our understanding and also revealed much about the structure of optimal controllers in these special cases where linear controllers are optimal. However, all of these quadratic-invariance structured problems also have no signaling incentive and the information patterns are nested [83].[1]

On the other hand, the second direction studies general cases when linear strategies are not optimal. Nayyar *et al.* discussed the structure of the optimal controllers in general decentralized problems [72], and Wu *et al.* found the mathematical properties (like continuity of the optimal strategy) of the optimal strategy for Witsenhausen's counterexample [110]. However, these results do not give quantifiable results, and to get such results we have to study the effect of implicit communication [45, 108].

Most of quantifiable results focus on Witsenhausen's counterexample. As we can see in the second box of Fig. 4.1, in Witsenhausen's counterexample the two controllers act in different time slots and may try to communicate. Exploiting implicit communication between the controllers makes nonlinear strategies outperform linear ones. Mitter and Sahai found that linear strategies can be arbitrarily bad compared to nonlinear strategies [68]. Many researchers including [59, 7, 61, 48] tried using computer-based exhaustive search to find the optimal strategy. Finally, Grover *et al.* showed that signaling-based nonlinear strategies approximately achieve the optimal cost to within a constant ratio [37]. This chapter continues this approach, and can be considered as a direct

[1]There are a few special cases when a linear controller is optimal but cannot be explained in the context of signaling incentives. Especially, in [10], Bansal and Basar found that when input cost and state disturbance measures match, a linear controller is optimal. Likewise, in communication theory where the encoder and decoder can be thought of as distributed controllers, it is well known that linear is optimal when the source and channel distributions and cost measures match [34]. This intriguing link deserves further exploration in future work.

descendent of [37]. In fact, there is a close relationship between Witsenhausen's counterexample and the scalar infinite-horizon LQG problem considered in this chapter. We will revisit this point in Section 4.4, and see that the infinite-horizon LQG problem can be thought of as the interlocking of a series of generalized Witsenhausen's counterexamples.

Another not directly but conceptually related branch of the second direction is "Control over Communication Channels" [97, 86, 119, 26, 69, 66], which tries to quantify explicit information flow for control. They introduce an explicit communication link and measure the amount of information flow required to control the system. One of their main results is that in scalar systems we need at least the communication rate, (log of eigenvalue) bits, to stabilize the system [97]. Later, this concept was extended to nonlinear filtering [64]. In this chapter, we will see the underlying relationship to decentralized control problems.

On the other hand, communication theory (especially, wireless communication theory) has developed a lot of quantifiable results for network communication problems. Since communication problems are decentralized in nature, the exact characterization of the capacity has been open for most communication networks involving many nodes. However, they still made progress by dividing cases based on the SNR (Signal-to-Noise Ratio), bringing linear-algebraic ideas and concepts to problems, and solving problems approximately. Especially, Avestimehr *et al.* considered relay communication problems with arbitrarily large number of nodes, and successfully characterize the capacity to within a constant gap that only scales with network size. At the heart of this progress, there are the concepts of *generalized degree of freedom (d.o.f.)* and *binary deterministic models*. In [6], Avestimehr *et al.* idealized bit levels as different antennas. By conceptualizing each bit level as a different subspace, they could apply linear-algebraic concepts and ideas for much precise analysis. By expanding the concept of d.o.f. (essentially, the rank of linear spaces) to different bit levels, they could understand the capacity of wireless communication networks to within a constant gap.

The main contribution of this chapter is the parallelism between information flows in decentralized LQG control and those in wireless communication theory. We will see that just as wireless communication theory divides cases depending on the SNR, decentralized LQG problems can be divided based on the eigenvalue of the systems. Moreover, we will find the relevant bottleneck in decentralized LQG problems using the idea of 'geometric slicing', which we believe is a proper analogy to the information-theoretic cutset bound [21] in a dynamic-programming context.

The rest of the chapter is organized as follows: We formally state the problem and the main results in Section 4.2. Section 4.3 gives the underlying intuitions behind the results. In Section 4.4, 4.5, 4.6, 4.7, we will convert these intuitions into formal proofs, and introduce proof ideas for that. Section 4.8 discusses the fundamental relationship between wireless communication theory and decentralized LQG problems. Finally, Section 4.9 concludes the chapter.

4.2 Problem Statement and Main Results

Throughout this chapter, we will discuss the scalar infinite-horizon decentralized LQG problems with two controllers.

Problem A (scalar infinite-horizon decentralized LQG problems with two controllers)**.**

$$x[n+1] = ax[n] + b_1u_1[n] + b_2u_2[n] + w[n]$$
$$y_1[n] = c_1x[n] + v_1[n]$$
$$y_2[n] = c_2x[n] + v_2[n]$$

Here, $u_1[n]$ and $u_2[n]$ must be causal functions of $y_1[n]$ and $y_2[n]$ respectively, i.e. $u_1[n] = f_{1,n}(y_1[0], \cdots, y_1[n])$ and $u_2[n] = f_{2,n}(y_2[0], \cdots, y_2[n])$. Following the traditional LQG problem formulation, the objective is minimizing an average quadratic cost:

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{0\le n<N} q\mathbb{E}[x^2[n]] + r_1\mathbb{E}[u_1^2[n]] + r_2\mathbb{E}[u_2^2[n]]. \tag{4.1}$$

Here, $q \ge 0$, $r_1 \ge 0$, $r_2 \ge 0$ and the underlying random variables are independent Gaussian, i.e. $x[0] \sim \mathcal{N}(0, \sigma_0^2)$, $w[n] \sim \mathcal{N}(0, \sigma_w^2)$, $v_1[n] \sim \mathcal{N}(0, \sigma_{v1}^2)$ and $v_2[n] \sim \mathcal{N}(0, \sigma_{v2}^2)$.

Figure 4.1 shows a pictorial description of the problem by introducing duplicated nodes across different time-steps and thus unraveling the dynamics.[2] First, without loss of generality, we put a series of assumptions on the problems.

Assumption (a): $b_1 = b_2 = 1$.

Assumption (b): $c_1 = c_2 = 1$.

Assumption (c): $\sigma_w^2 = 1$.

Assumption (d): $\sigma_{v1} \le \sigma_{v2}$.

Assumptions (a), (b) do not lose generality since we can rescale u_1, u_2 and y_1, y_2 respectively. Assumption (c) doesn't lose generality since we can rescale the system equation by $\frac{1}{\sigma_w}$. Assumption (d) doesn't lose generality because it is simply a way of deciding which controller is 1, and which is 2.

Therefore, throughout this chapter we will consider the following problem:

Problem B (Normalized decentralized LQG problem for Problem A)**.**

$$x[n+1] = ax[n] + u_1[n] + u_2[n] + w[n] \tag{4.2}$$
$$y_1[n] = x[n] + v_1[n]$$
$$y_2[n] = x[n] + v_2[n]$$

where $x[0] \sim \mathcal{N}(0, \sigma_0^2)$, $w[n] \sim \mathcal{N}(0, 1)$, $v_1[n] \sim \mathcal{N}(0, \sigma_{v1}^2)$, $v_2[n] \sim \mathcal{N}(0, \sigma_{v2}^2)$. The control objective is minimizing the long-term average cost in (4.1)*.*

[2] The idea of unraveling the system by introducing duplicated nodes across different time-steps was also used to study network information flows [1]. As in [1], we will see the unraveling of the dynamics will be helpful to find the information bottleneck of the system.

Even though this problem is the simplest decentralized infinite-horizon LQG problem, as we will see in Proposition 4.4, linear strategies are not optimal and the optimization problem becomes infinite-dimensional and non-convex. Here, we follow the approximation approach of [37], which itself inherits from [6] and the spirit of approximate algorithms in computer science theory. We propose a set of finite-dimensional function spaces that are guaranteed to contain an approximately optimal solution. Therefore, if we optimize only over the proposed finite-dimensional function spaces, the solution achieves the optimal performance within a constant ratio regardless of the problem parameters, a, q, r_1, r_2, σ_0, σ_{v1}, and σ_{v2}. In this chapter, we first consider the fast-dynamics case when the single eigenvalue of the system is large ($|a| \geq 2.5$) and discuss the conceptual relationship with high-SNR in wireless communication theory. The slow-dynamics case when the single eigenvalue of the system is small ($|a| < 2.5$) will be discussed in Chapter 5, and the relationship with low-SNR in wireless communication theory will be also revealed.[3]

The first set of controllers is two naive memoryless linear strategies, which simply zero-force the state.

Definition 4.1 (Linear Strategy $L_{lin,bb}$)**.** *$L_{lin,bb}$ is the set of the following two controllers:*
(i) $u_1[n] = -ay_1[n]$, $u_2[n] = 0$.
(ii) $u_1[n] = 0$, $u_2[n] = -ay_2[n]$.

The second set is a two-parameter (s, d) nonlinear strategy set for implicit communication between two controllers.

Definition 4.2 (s-Stage Signaling Strategy $L_{sig,s}$)**.** *For a given $s \in \mathbb{N}$, $L_{sig,s}$ is the set of all controllers which can be written as the following form for some $d > 0$,*

$$u_1[n] = -aR_d(y_1[n]) \tag{4.3}$$

$$u_2[n] = -a(Q_{a^s d}(y_2[n] - R_{a^s d}(\sum_{1 \leq i \leq s} a^{i-1}u_2[n-i])) + R_{a^s d}(\sum_{1 \leq i \leq s} a^{i-1}u_2[n-i])). \tag{4.4}$$

Here, $Q_x(y) := x\lfloor \frac{y}{x} + \frac{1}{2} \rfloor$ and $R_x(y) := y - Q_x(y)$. These quantities represent the quantization level and remainder when y is divided by x respectively, i.e. let $y = q \cdot x + r$ for $q \in \mathbb{Z}$ and $r \in [-\frac{x}{2}, \frac{x}{2})$. Then, $Q_x(y) = q \cdot x$ and $R_x(y) = r$. (We also put $u_1[n] = 0$ and $u_2[n] = 0$ for $n < 0$.)

We will give the intuition behind this strategy in Section 4.3. Roughly speaking, in the strategy set $L_{sig,s}$ the first controller "implicitly communicates" its observation to the second controller with delay s by making the state easier to estimate. This strategy is essentially a multi-stage generalization of the lattice-quantization strategy [37] used for Witsenhausen's counterexample. Notice that the strategy requires remembering the past s control inputs.

[3] Here, we did not optimize for the best ratio, and the explicit number 2.5 is arbitrary. We could have written Theorem 4.1, 4.2 with any fixed number like $|a| = 2, 3, 5, 6, \cdots$ which may result in a different ratio. However, as $|a|$ increases, the ratio between linear and nonlinear strategy cost goes to infinity, and the gain by considering nonlinear strategies becomes larger.

Now, we can state the main theorem of this chapter, which tells us that when $|a| \geq 2.5$ optimizing over $L_{lin,bb}$ and $L_{sig,s}$ is enough to give a constant-ratio optimal strategy among all possible strategies.

Theorem 4.1. *Consider the decentralized LQG problem shown in Problem B. Let $L' = L_{lin,bb} \cup \bigcup_{s\in\mathbb{N}} L_{sig,s}$ and L be the set of all measurable causal strategies. There exists a constant $c \leq 1.5\times 10^5$ such that for all $|a| \geq 2.5$, q, r_1, r_2, σ_0, σ_{v1} and σ_{v2},*

$$\frac{\inf_{u_1,u_2\in L'} \limsup_{N\to\infty} \frac{1}{N}\sum_{0\leq n<N} \mathbb{E}[qx^2[n] + r_1u_1^2[n] + r_2u_2^2[n]]}{\inf_{u_1,u_2\in L} \limsup_{N\to\infty} \frac{1}{N}\sum_{0\leq n<N} \mathbb{E}[qx^2[n] + r_1u_1^2[n] + r_2u_2^2[n]]} \leq c.$$

Proof. See Section 4.7. □

Since measurability is the minimal condition required to even have the problem make any sense, $\inf_{u_1,u_2\in L}$ implies a minimization over all possible control strategies. Thus, in the rest of the chapter, we will simply write it as $\inf_{u_1,u_2}$.

For the proof, we give explicit and computable upper and lower bounds on the optimal cost, and prove that they are within a constant ratio. In Lemma 4.15 of page 214, we will see that the linear strategies $L_{lin,bb}$ give the following upper bounds on the minimal average cost.

$$\inf_{u_1,u_2} \limsup_{N\to\infty} \frac{1}{N}\sum_{0\leq n<N} \mathbb{E}[qx^2[n] + r_1u_1^2[n] + r_2u_2^2[n]] \leq q(a^2\sigma_{v1}^2+1) + r_1(a^4\sigma_{v1}^2 + a^2\sigma_{v1}^2 + a^2).$$

$$\inf_{u_1,u_2} \limsup_{N\to\infty} \frac{1}{N}\sum_{0\leq n<N} \mathbb{E}[qx^2[n] + r_1u_1^2[n] + r_2u_2^2[n]] \leq q(a^2\sigma_{v2}^2+1) + r_2(a^4\sigma_{v2}^2 + a^2\sigma_{v2}^2 + a^2).$$

In Lemma 4.7 of page 174, we will see that the signaling strategies $L_{sig,s}$ give the following upper bounds.

$$\begin{aligned}&\inf_{u_1,u_2} \limsup_{N\to\infty} \frac{1}{N}\sum_{0\leq n<N} \mathbb{E}[qx^2[n] + r_1u_1^2[n] + r_2u_2^2[n]]\\ &\leq \inf_{(d,w_1)\in S_{U,1}} qD_{U,1}(d,w_1) + r_1\frac{a^2d^2}{4} + r_2(8a^2D_{U,1}(d,w_1) + \frac{7}{2}a^{2(s+1)}d^2 + 4a^2\sigma_{v2}^2).\end{aligned}$$

where the definitions of $D_{U,1}(d,w_1)$ and $S_{U,1}$ are available in Lemma 4.7 of page 174.

For the lower bounds, we will see four different bounds in Lemma 4.12 of page 198 and Lemma 4.13 of page 208. Thus, the optimal cost of Problem B is lower bounded as follows.

$$\begin{aligned}&\inf_{u_1,u_2} \limsup_{N\to\infty} \frac{1}{N}\sum_{0\leq n<N} q\mathbb{E}[x^2[n]] + r_1\mathbb{E}[u_1^2[n]] + r_2\mathbb{E}[u_2^2[n]]\\ &\geq \max_{1\leq i\leq 4} \sup_{(k_1,k_2,k,\sigma'_{v2},\alpha,\Sigma)\in S_{L,i}} \inf_{\widetilde{P_1},\widetilde{P_2}\geq 0} qD_{L,i}(\widetilde{P_1},\widetilde{P_2};k_1,k_2,k,\sigma'_{v2},\alpha,\Sigma) + r_1\widetilde{P_1} + r_2\widetilde{P_2}.\end{aligned}$$

Here, the definitions of $D_{L,1}$ and $S_{L,1}$ are available in Lemma 4.12 of page 198. The remaining definitions of $D_{L,i}$ and $S_{L,i}$ for $2 \leq i \leq 4$ are shown in Lemma 4.13 of page 208.

Finally, in Section 4.7 of page 212 we will compare these upper and lower bounds, and prove that they are within a constant ratio.

To prove a similar result for the slow-dynamics case ($|a| < 2.5$), we need a further set of single-controller optimal strategies. These strategies are linear strategies which can be parametrized by a single parameter k.

Definition 4.3. *$L_{lin,kal}$ is the set of all controllers which can be written in either one of the two following forms for some $k \in \mathbb{R}$*
(i) $u_1[n] = -k\mathbb{E}[x[n]|y_1^n, u_1^{n-1}]$, $u_2[n] = 0$.
(ii) $u_1[n] = 0$, $u_2[n] = -k\mathbb{E}[x[n]|y_2^n, u_2^{n-1}]$.

Here, $\mathbb{E}[x[n]|y_1^n, u_1^{n-1}]$ and $\mathbb{E}[x[n]|y_2^n, u_2^{n-1}]$ can be easily computed by Kalman filtering once k is fixed.[4]

The results of Chapter 5 will show that when $|a| \leq 2.5$, optimization over $L_{lin,kal}$ is enough to give a constant-ratio optimal strategy among all possible strategies.

Theorem 4.2. *There exists $c \leq 2 \cdot 10^6$ such that for all $|a| \leq 2.5$, q, r_1, r_2, σ_0, σ_{v1} and σ_{v2},*

$$\frac{\inf_{u_1,u_2 \in L_{lin,kal}} \limsup_{N\to\infty} \frac{1}{N} \sum_{0\leq n<N} \mathbb{E}[qx^2[n] + r_1u_1^2[n] + r_2u_2^2[n]]}{\inf_{u_1,u_2} \limsup_{N\to\infty} \frac{1}{N} \sum_{0\leq n<N} \mathbb{E}[qx^2[n] + r_1u_1^2[n] + r_2u_2^2[n]]} \leq c.$$

Proof. This will be shown in Theorem 5.1 of Chapter 5. □

By Theorem 4.1 and 4.2, we can achieve the optimal cost to within a factor of $2 \cdot 10^6$ by optimizing only over $L_{lin,kal}$ and $L_{sig,s}$, which only involves single and two parameter optimization problems respectively. We believe that the factor here is coming from our proof strategy and the gap will be far smaller in practice.

The optimal parameters for the proposed strategy sets in Definition 4.1, 4.2, 4.3 are not difficult to find. The optimization over $L_{lin,kal}$ is a centralized LQG problem and it is well known that the optimal k can be easily found by Riccati equations [55]. In Proposition 4.7 of page 215 and Proposition 4.8 of page 216, we will see that the parameter s in $L_{sig,s}$ can be selected based on the problem parameters. Particularly, we can use $s = \lceil \frac{\ln \sigma_{v2}^2 - \ln(\max(1, a^2\sigma_{v1}^2))}{2\ln a} \rceil$, so that $a^{2(s-1)} \max(1, a^2\sigma_{v1}^2) < \sigma_{v2}^2 \leq a^{2s} \max(1, a^2\sigma_{v1}^2)$. Moreover, Corollary 4.3 of page 216 gives a simple analytic upper bound on the performance of $L_{sig,s}$, which has only two local optima as d varies. Therefore, both optimization problems are easily solvable.

However, the true implication of Theorem 4.1 and 4.2 is that they reveal the essential skeletons of an optimal strategy. Since the original optimization problems are infinite-dimensional

[4] Since u_i^{n-1} is known to the controller, we can compensate for the past control inputs and treat it as an open-loop system. The estimation problem in the open-loop system is well-known Kalman filtering. This concept is called the control-estimation separation principle in the control community.

and non-convex, it is not even clear how and where to start a computer-based search. By revealing the minimal strategy for approximately optimal performance, these results might give an initial point to start optimization for further performance refinements. More importantly, as we will see in later sections, the proposed strategy sets are intuitively interpretable and understandable.

4.3 Intuition: Deterministic Model Interpretation

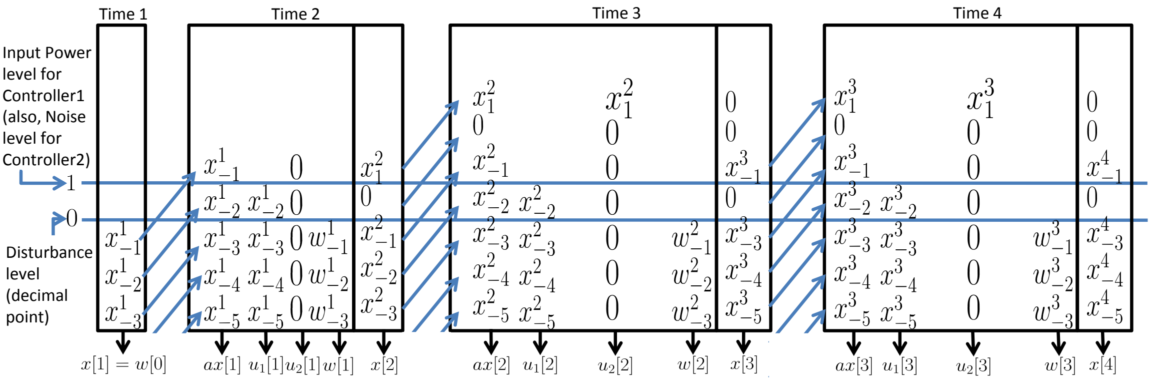


Figure 4.2: Deterministic Model Interpretation of Nonlinear Control Strategies $L_{sig,1}$

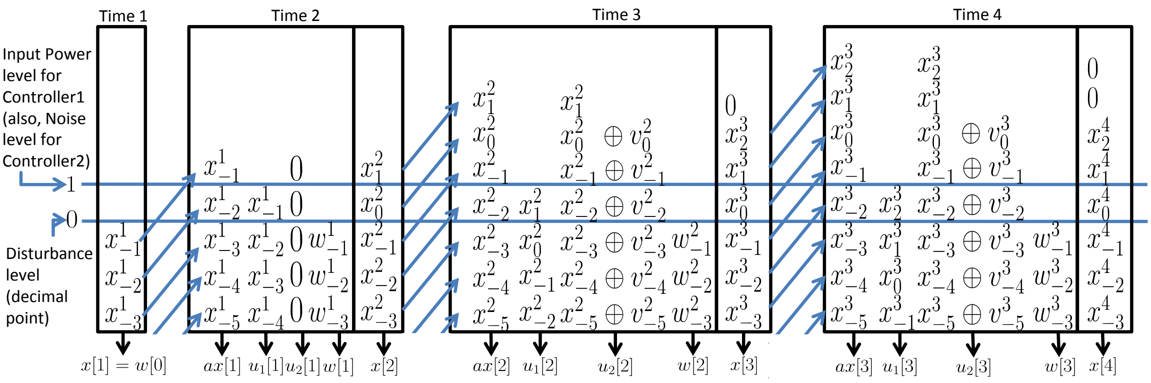


Figure 4.3: Deterministic Model Interpretation of Linear Control Strategies L'_{lin}

After reading the problem statement and the main result, readers may wonder

(1) Why are linear strategies not enough to achieve a constant ratio from the optimal?

(2) Why is the proposed set of nonlinear strategies enough to achieve a constant ratio from the optimal cost?

In this section, we will give an intuitive answer for these questions based on a linear deterministic model in the spirit of Avestimehr, Diggavi and Tse [6], which has already proved to be useful in understanding some control problems [37, 81].

The point of these linear deterministic models is to simplify and idealize real arithmetic, and allow us to take a linear view of nonlinearity. The idea is to consider real numbers in binary expansion and then to simplify arithmetic by eliminating carries. For example, if we have a number 5 we write it as 101 in binary expansion, and likewise we write $\frac{1}{4}$ as 0.01. If we have a random variable X which is uniform on $[0, 4)$, we can write it as $b_1 b_0 . b_{-1} b_{-2} \cdots$ in binary expansion where b_i are i.i.d. Bernoulli($\frac{1}{2}$) random variables on $\{0, 1\}$.

Since uniform random variables are so simple, we idealize Gaussian random variables as uniform random variables. For a given Gaussian random variable with zero mean and variance σ^2, we caricature it as a uniform random variable on $[0, \sigma)$ and use the same deterministic model for the uniform random variable. For example, a Gaussian random variable $\mathcal{N}(0, 4^2)$ is caricatured as $b_1 b_0 . b_{-1} b_{-2} \cdots$.

Then, we simplify the arithmetic on binary representations. Addition and subtraction are approximated by bitwise XOR — thereby ignoring the carry effect. For $B' = b'_1 b'_0 . b'_{-1} b'_{-2} \cdots$ and $B'' = b''_1 b''_0 . b''_{-1} b''_{-2} \cdots$, both $B' + B''$ and $B' - B''$ are approximated by $(b'_1 \oplus b''_1)(b'_0 \oplus b''_0).(b'_{-1} \oplus b''_{-1})(b'_{-2} \oplus b''_{-2}) \cdots$. Since we are modeling addition and subtraction in the same way, we ignore the sign of the numbers and consider x and $-x$ to be the same in deterministic models and so we will assume every number is positive from now on.

Multiplication is approximated by a bit shift. For example, $B' \times 4$ and $B'/4$ are equal to $b'_1 b'_0 b'_{-1} b'_{-2} . b'_{-3} \cdots$ and $0.b'_1 b'_0 b'_{-1} \cdots$ respectively. If we restrict multipliers and dividers to be 2^n, this agrees with conventional multiplication and division.

For further discussion, it will be helpful to define the (binary) index and level for binary expansions. For a given binary expansion $B = \cdots b_1 b_0 . b_{-1} b_{-2}$, the *index* i bit of B indicates b_i. It is natural to call the bits $b_i, b_{i+1}, b_{i+2}, \cdots$ as the bits above level i, and the bits $b_{i-1}, b_{i-2}, b_{i-3}, \cdots$ as the bits below level i. To clarify this point, we define the *level* i as the imaginary line between two sequential bits b_i and b_{i-1}. Thus, the decimal point corresponds to the level 0.

We also denote the *upper-level* of B as the minimum level l such that all bits above the level l are 0, i.e. $b_i = 0$ for all $i \geq l$. For example, the upper level of 3 is 2 and the upper level of 4 is 3. When B is a random variable, we denote the *upper level* l of B as the worst case bound, i.e. the minimum l such that $b_i = 0$ for all $i \geq l$ with probability 1. Therefore, the upper level of a uniform random variable on $[0, 4)$ is 2 since 4 is not included in the interval.

Now, we can come up with the corresponding binary deterministic counterparts for the LQG problems of Problem B. To simplify the discussion, we will assume the first controller has perfect observations, and the second controller has no input cost. Like [37], we will consider the state minimization problem for given control power constraints, instead of the weighted long-term cost minimization problem.

Problem C (Binary Deterministic Model for Problem B)**.** *Let $a', \sigma'_{v2}, p'_1 \in \mathbb{Z}$ be the problem parameters. For time index $n \geq 0$ and binary index i, the deterministic system dynamics is given as follows:*

$$x_i^{n+1} = x_{i-a'}^n \oplus u_{1,i}^n \oplus u_{2,i}^n \oplus w_i^n$$
$$y_{1,i}^n = x_i^n$$
$$y_{2,i}^n = x_i^n \oplus v_i^n$$

Here, $x_i^0 = 0$ for all i. For all n, w_i^n are 0 for all $i \geq 0$ and i.i.d. Bernoulli $\frac{1}{2}$ on $\{0,1\}$ for all $i < 0$. For all n, v_i^n are 0 for all $i \geq \sigma'_{v2}$ and i.i.d. Bernoulli $\frac{1}{2}$ on $\{0,1\}$ for all $i < \sigma'_{v2}$. The v_i^n are independent from the w_i^n. $u_{1,i}^n$ and $u_{2,i}^n$ are causal functions of $y_{1,i}^n$ and $y_{2,i}^n$ respectively. The first controller has a "power" limit, $u_{1,i}^n = 0$ for all $n \geq 0$ and $i \geq p'_1$.

The goal of control is to minimize the upper-level d on the worst state distortion, i.e. minimizing d such that $x_i^n = 0$ for all $i \geq d$ and $n \geq 0$ with probability 1.

We can notice that x_i^n, $u_{1,i}^n$, $u_{1,i}^n$, w_i^n, v_i^n correspond to $x[n]$, $u_1[n]$, $u_2[n]$, $w[n]$, $v[n]$ of Problem B respectively. Therefore, we will use the latter terms for a compact representation of the bits in Problem C. Moreover, since the parameters of Problem C are given in the binary levels of amplitude, they have the following relationship with those of Problem B: $a = 2^{a'}$, $\sigma_{v2}^2 = 2^{2\sigma'_{v2}}$, $\mathbb{E}[u_1^2[n]] \leq 2^{2p'_1}$. Through the rest of discussion, we will focus on the case, $a' = 2, \sigma'_{v2} = \frac{a'}{2}, p'_1 = \frac{a'}{2}$. Therefore, the corresponding parameters in LQG Problem B are $a = 2^2$, $\sigma_w^2 = 1$, $\sigma_{v2}^2 = 2^2$, $\mathbb{E}[u_1^2[n]] \leq 2^2$.

Based on this deterministic model, we will answer the first question, 'why the proposed strategy is approximately optimal'. First, we can easily derive the following lower bound on the state disturbance.

Proposition 4.1. *When $a' = 2, \sigma'_{v2} = \frac{a'}{2}, p'_1 = \frac{a'}{2}$ in Problem C, the minimum upper-level d on the state distortion level has to be at least 2.*

Proof. We can easily see that for $n \geq 1$, the distortion level of $x[n]$ is at least 0 since $w[n-1]$ with upper-level 0 is added at each time step. At time $n+1$, this distortion will be shifted up by two bits, and the upper-level of the distortion becomes 2. However, the first controller cannot touch the bits above the level 1 and so the first controller cannot reduce the distortion level. Moreover, at the second controller, any bits below level 1 are corrupted by i.i.d. Bernoulli($\frac{1}{2}$) observation noise. Therefore, the second controller's observation is independent from the unknown bits sitting below the level 1. Consequently, there is no action it can take to draw that bit to 0.

Neither controller can reduce the distortion bit sitting between the level 1 and 2, so with a positive probability $\frac{1}{2}$ this bit can be non-zero. Therefore, the upper-level of $x[n+1]$ has to be at least 2, i.e. $d \geq 2$. □

In fact, the following proposition shows this lower bound is actually achievable.

Proposition 4.2. *Consider Problem C with* $a' = 2, \sigma'_{v2} = \frac{a'}{2}, p'_1 = \frac{a'}{2}$*. Given the first controller's observation* $y_1[n] = \cdots y^n_{1,1} y^n_{1,0}.y^n_{1,-1}\cdots$ *and the second controller's observation* $y_2[n] = \cdots y^n_{2,1} y^n_{2,0}.y^n_{2,-1}\cdots$*, let the first and second controller's control input be*

$$u_1[n] = y^n_{1,-2}.y^n_{1,-3}y^n_{1,-4}\cdots$$
$$u_2[n] = \cdots y^n_{2,2}y^n_{2,1}000.0\cdots$$

Then, this strategy can achieve the optimal upper-level on the state distortion, $d = 2$*.*

Proof. Since we already know the minimal $d \geq 2$ from Proposition 4.1, it is enough to show that the proposed strategy can achieve $d = 2$.

Figure 4.2 shows the resulting dynamics when we actually use this strategy. Since the initial state $x[0] = 0$, at time 1 both controllers' inputs are also 0 and $w[0]$ is only term that contributes to $x[1]$. Thus, $x[1]$ can be represented by $0.x^1_{-1}x^1_{-2}x^1_{-3}\cdots$ in the deterministic model where each bit is i.i.d. Bernoulli $\frac{1}{2}$ in $\{0, 1\}$.

At time 2, $x[1]$ is shifted two-bits up to generate $x^1_{-1}x^1_{-2}.x^1_{-3}\cdots$. Since the first controller's observation $y_1[1]$ is equal to $x_1[n]$, its control input is $x^1_{-2}.x^1_{-3}\cdots$. After being corrupted by the noise $v_2[1]$, the second controller's observation $y_2[1]$ becomes $(v^1_0).(x^1_{-1} \oplus v^1_{-1})(x^1_{-2} \oplus v^1_{-2})\cdots$ which is independent from the state, and as a result $u_2[1] = 0$. When all of these are added, the second bit of the state canceled by the first controller's input. As we can see in Figure 4.2 the state $x[2]$ results in $x^2_1 0.x^2_{-1}x^2_{-2}\cdots$ where each bit is i.i.d. Bernoulli $\frac{1}{2}$ except for the 0 in the 0th position.

At time 3, the first controller does essentially the same operation as time 2, canceling the lower bits below the level 1. However, the second controller's observation has larger level than before, $y_2[2] = (x^2_1)(v^2_0).(x^2_{-1} \oplus v^2_{-1})\cdots$. Thus, $u_2[2]$ becomes $x^2_1 000.0\cdots$. When we are adding these values, the first bit of the state is canceled by the second controller's input and the third bit of the state is canceled by the first controller's input. As Figure 4.2 shows, the resulting state $x[3]$ is $x^3_1 0.x^3_{-1}x^3_{-2}\cdots$, which is essentially the same as $x[2]$.

Therefore, we arrive in steady state and repeating the control strategy always gives the state with the same upper-level 2. This finishes the proof. □

So, we have an optimal scheme for the deterministic model. Let's apply the insights that we learnt from the deterministic model to the original LQG problem. The first controller's strategy of Proposition 4.2 can be understood as a sequence of two operations. The first operation is extracting the lower bits of $y_1[n]$ and thus generating $0.0y_{1,-2}y_{1,-3}\cdots$. To mimic this, we can simply divide $y_1[2]$ by 0.1 (in binary) and take the remainder. Using Definition 4.2, this can be written as $R_d(y_1[2])$ with $d = 0.1$. The second operation is shifting the bits up to generate $u_1[2]$, which is just multiplication by a constant ($-a$ to be exact). Therefore, $u_1[2] = -aR_d(y_1[2])$.

The second controller's strategy of Proposition 4.2 can be understood as a sequence of two operations. The first operation is extracting the higher bits of $y_2[2]$ and thus generating $y_{2,m}\cdots y_{2,1}0.0\cdots$. For this, we can divide $y_2[2]$ by 10 (in binary) and take the quotient (exactly speaking, quotient multiplied by divisor). Using Definition 4.2, this can be written as $Q_{d'}(y_2[2])$ with $d' = ad = 10$. The second operation is shifting two bits up as before, which is multiplication. Therefore, $u_2[2] = -aQ_{d'}(y_2[2])$.

Compared to (4.3) and (4.4), this strategy is essentially equivalent to 1-stage signaling strategy except for some minor terms in $u_2[n]$. Therefore, in nonlinear strategies $L_{sig,s}$, $u_1[n]$ tries to cancel the lower bits in $ax[n]$ by exploiting its better observation and $u_2[n]$ tries to cancel higher bits in $ax[n]$ exploiting its less expensive input cost.

Now, we understand why the proposed strategy might be approximately optimal. We can move on to the next question, 'why linear is not enough for constant-ratio optimality'. Let's first remind ourselves of the counterparts to linear operations in these binary models. Addition and subtraction correspond to bitwise XOR. Multiplication and division by a constant correspond to shifting bits up and down. Let's revisit Problem C with these restrictions on the strategy, and understand why we cannot achieve the optimal performance with linear strategies.

Proposition 4.3. *Consider Problem C with $a' = 2, \sigma'_{v2} = \frac{a'}{2}, p'_1 = \frac{a'}{2}$. Let's restrict the controller strategies to the following forms: For some $k_{i,j}, k'_{i,j} \in \mathbb{Z}$ and for all $i \in \mathbb{Z}$ and $n \in \mathbb{Z}^+$,*

$$u^n_{1,i} = y^0_{1,i+k_{0,n}} \oplus y^1_{1,i+k_{1,n}} \oplus \cdots \oplus y^n_{1,i+k_{n,n}}$$
$$u^n_{2,i} = y^0_{2,i+k'_{0,n}} \oplus y^1_{2,i+k'_{1,n}} \oplus \cdots \oplus y^n_{2,i+k'_{n,n}}$$

Under this constraint on control strategy, the minimal upper-level d on the state distortion is 3.

Intuitively this proposition is obvious. As we saw in Proposition 4.1, the first controller has input power is strictly less than the distortion level 2. When we restrict the strategy to be linear, the first controller cannot cancel any bits in the state. Therefore, the second controller is the only controller that can control the state. The second controller can only see the bits above level 1, and after one time step, the distortion level will become 3. Let's clarify this point more carefully.

Time 1 is the same as the proof of Proposition 4.2. However, at time 2 the first controller cannot cancel the lower bits any more. The only allowed operations are shifting the bits in each observation and taking XOR between them. As we can see in Figure 4.3, within the power constraint the first controller cannot but shift at least one-level down the bits in $y_1[1]$, and may choose $u_1[1] = x^1_{-1}.x^1_{-2}x^1_{-3}\cdots$. As we discussed in Proposition 4.2, the second controller's observation is independent from the state and the optimal $u_2[1]$ is 0. Therefore, as we can see in Figure 4.3 no bits cancel with each other, and $x[2] = x^2_1x^2_0.x^2_{-1}\cdots$ where each bits are i.i.d. Bernoulli $\frac{1}{2}$.

At time 3, due to the same reason, the best feasible input for the first controller is $u_1[2] = x^2_1.x^2_0x^2_{-1}\cdots$ and cannot cancel any bits in the state. Meanwhile, the second controller's observation with additive noise is $y_2[2] = x^2_1(x^2_0 \oplus v^2_0).(x^2_{-1} \oplus v^2_{-1})\cdots$. Therefore, to cancel the first bits of the state, the second controller shifts two-bits up in $y_2[2]$ and chooses $u_2[2] = x^2_1\ (x^2_0 \oplus v^2_0)\ (x^2_{-1} \oplus v^2_{-1})\ (x^2_{-2} \oplus v^2_{-2}).\ (x^2_{-3} \oplus v^2_{-3})\cdots$. When these are added, the first bit of the state cancels and the resulting state $x[3]$ has three bits above the decimal point as we can see in Figure 4.3.

By repeating this procedure, we can see that after the transient states of time steps $1, 2, 3$, the plant and controllers stay in steady state. From Figure 4.3, in steady state, $x[n]$ has three bits above the decimal point. Therefore, compared to the optimal performance without the linear controller constraint, we can see one-bit-level performance degradation. This degradation comes from the inefficient use of the first controller input. In other words, the first controller cancels the lower bits of the state in the optimal strategy while it cannot cancel any bits in linear strategies.

Let's consider the Gaussian counterpart of the previous results. As we discussed earlier, the corresponding parameters in the original LQG problem is $\sigma_{v1}^2 = 0$, $\sigma_{v2}^2 = \Theta(a)$, $\mathbb{E}[u_1^2[n]] \leq \Theta(a)$, $\mathbb{E}[u_2^2[n]] \leq \infty$. We will consider the minimum state distortion as a goes to infinity. From Proposition 4.2, we can expect that the optimal state distortion is $\mathbb{E}[x^2[n]] \leq O(a^2)$ with these parameters.[5] From Proposition 4.3, we can expect that the state distortion is $\mathbb{E}[x^2[n]] \geq \Omega(a^3)$ when we restrict control strategies to be linear. Here, we can see the ratio between the optimal cost and the linear strategy cost goes to infinity as a grows.

Even if the discussion so far focused on minimizing the state distortion under power constraints, the result can be easily converted to the weighted long-term cost. Let's choose the parameters of Problem B as $q = 1$, $r_1 = a$, $r_2 = 0$, $\sigma_0 = 0$, $\sigma_{v1}^2 = 0$, and $\sigma_{v2}^2 = a$, i.e.

$$\inf_{u_1,u_2} \limsup_{N\to\infty} \frac{1}{N} \sum_{0\leq n<N} \mathbb{E}[x^2[n] + a u_1^2[n]]$$

If $\mathbb{E}[x^2[n]] \leq O(a^2)$ when $\mathbb{E}[u_1^2[n]] \leq \Theta(a)$ as we predicted, the optimal weighted cost has to be $O(a^2)$. However, if we restrict the control strategies to be linear, $\mathbb{E}[x^2[n]]$ will be $\Omega(a^3)$ with the same power constraint according to our conjecture. Therefore, we need at least $\mathbb{E}[u_1^2[n]] \geq \Theta(a^2)$ to make $\mathbb{E}[x^2[n]] \leq O(a^2)$. In either case[6], the weighted cost is $\Omega(a^3)$.

Formally, the following proposition formalizes this insight and proves the ratio between the optimal cost and linear strategy cost actually diverges in Gaussian problems.

Proposition 4.4. *Let L'_{lin} be the set of all linear time-varying controllers which can be written in the following form:*

$$u_1[n] = \sum_{i\leq n} k_{n,i} y_1[i],$$
$$u_2[n] = \sum_{i\leq n} k'_{n,i} y_2[i].$$

[5]Exactly speaking, the optimal state distortion is $\mathbb{E}[x^2[n]] \leq O(a^2 \log a)$ with input power constraint $\mathbb{E}[u_1^2[n]] \geq \Theta(a \log a)$. The is due to the fact that unlike uniform random variables Gaussian random variables can be arbitrarily large with exponentially decreasing probability. Later, this effect will be captured by large deviation ideas, and turns out to be crucial to get constant-ratio optimality. We will discuss more about this issue in Section 4.3.1.

[6]One may wonder why we do not consider the cases between $\mathbb{E}[u_1^2[n]] = \Theta(a)$ and $\mathbb{E}[u_1^2[n]] = \Theta(a^2)$, for example $\mathbb{E}[u_1^2[n]] = \Theta(a^{\frac{3}{2}})$. The reason comes from the limitation of these bit-wise deterministic models, precision. For $a = 4$, we will write $u_1[n]$ as $u_0^n.u_{-1}^n u_{-2}^n \cdots$ in binary when $\mathbb{E}[u_1^2[n]] = a$, and as $u_1^n u_0^n.u_{-1}^n \cdots$ when $\mathbb{E}[u_1^2[n]] = a^2$. When $\mathbb{E}[u_1^2[n]] = a^{\frac{3}{2}}$, we have to choose either one of these two. We choose the former in this chapter, so we cannot resolve the difference between $\mathbb{E}[u_1^2[n]] = a$ and $\mathbb{E}[u_1^2[n]] = a^{\frac{3}{2}}$.

Consider Problem B with parameters $q = 1$, $r_1 = a$, $r_2 = 0$, $\sigma_0^2 = 0$, $\sigma_{v1}^2 = 0$, *and* $\sigma_{v2}^2 = a$. *Then, we have*

$$\frac{\inf_{u_1,u_2 \in L'_{lin}} \limsup_{N \to \infty} \frac{1}{N} \sum_{0 \leq n < N} \mathbb{E}[qx^2[n] + r_1 u_1^2[n]]}{\inf_{u_1,u_2 \in L_{sig,1}} \limsup_{N \to \infty} \frac{1}{N} \sum_{0 \leq n < N} \mathbb{E}[qx^2[n] + r_1 u_1^2[n]]} \to \infty$$

as $a \to \infty$.

Proof. See Appendix 9.6. □

From the discussion above, we can see that the first controller with better observations is "signaling" to the second controller (with worse observations) through the control actions. However, the notion of communication here is different from the conventional one. In conventional communication problems, the transmitter has access to a source (but cannot change it) and reduces the uncertainty about the source at the destination by explicitly sending information about the source.

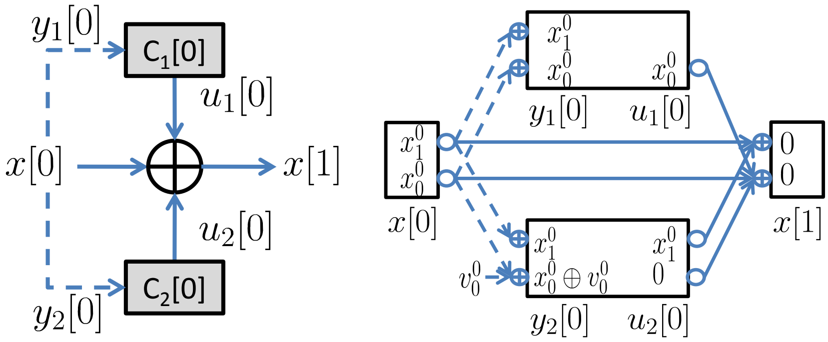
However, in control systems the source is the state, and the "transmitter"(which is a controller) can change the source itself by control action. Therefore, it can reduce the uncertainty of the source and make the source easier to estimate at the destination. Then, the destination will have a better idea about the source even without receiving any explicit information. This generalized notion of communication is the one happening between the first and the second controller.

Moreover, we can also see the delay of the communication is crucial in control problems, while this is usually ignored in traditional information theory. In Figure 4.2, the second controller has to wait until the disturbance is amplified above its observation noise level, which causes a 1-step delay between two controllers. However, as we increase the observation noise level of the second controller, the second controller has to wait longer until the disturbance is amplified enough and this will result in a longer "delay" between the two controller's actions.

In Section 4.5, we will explore this point by relating the infinite-horizon LQG problem to control problems with different time horizons. As we saw in Figure 4.1, Radner's problem [80] and Witsenhasuen's counterexample [108] are sub-blocks of the infinite-horizon LQG problem. We will see later in Section 4.5 that the scheme discussed here is a 1-step-delay implicit communication scheme which essentially (approximately) solves Witsenhausen's counterexample. In general, we may need up to an s-step-delay implicit communication to solve s-stage MIMO Witsenhausen's counterexamples.

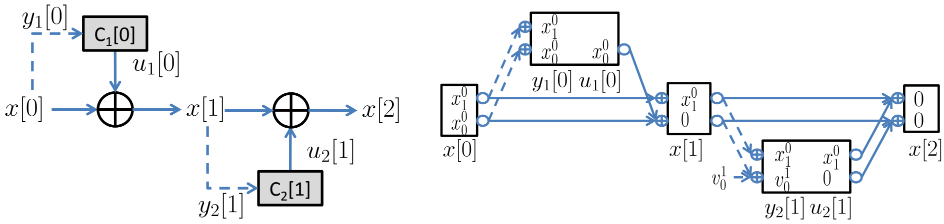
4.3.1 Caveat: Deterministic Model does not work for Radner's Problem

Even though we explained the result based on the binary deterministic model, it is just a simplified model for intuition and we should not naively believe that the same results always hold in Gaussian models as well. In fact, we will show that in Radner's problem [80] the deterministic model fails to correctly predict the behavior of Gaussian problems.



(a) (b)

Figure 4.4: (a) Radner's Problem and (b) the corresponding binary deterministic model. Here, the binary deterministic model can fail to correctly predict the optimal strategy and the optimal performance.



(a) (b)

Figure 4.5: (a) Witsenhausen's Counterexample and (b) the corresponding binary deterministic model . Here, the binary deterministic model does approximately predict the optimal strategy and the optimal performance.

In [80], Radner considered the following problem of Figure 4.4a: $x[0]$, $v_1[0]$, $v_2[0]$ are independent Gaussian random variables with zero mean and variance σ_0^2, σ_{v1}^2, σ_{v2}^2 respectively. Let $y_1[0] := x[0] + v_1[0]$, $y_2[0] := x[0] + v_2[0]$, $u_1[0] := f_1(y_1[0])$, $u_2[0] := f_2(y_2[0])$ and $x[1] := x[0] + u_1[0] + u_2[0]$. The control objective is minimizing $\mathbb{E}[qx[1]^2 + r_1u_1[0]^2 + r_2u_2[0]^2]$. And he proved that a linear controller is optimal.

Later, Witsenhausen found that if we shift the second controller by one time-step, the problem is fundamentally different and the optimal controller is not linear [108]. Figure 4.5a shows the counterexample: $x[0]$, $v_1[0]$, $y_1[0]$, $u_1[0]$ are the same as Radner's problem. However, $x[1] := x[0]+u_1[0]$, the second controller observes $y_2[1] := x[1]+v_2[1]$ where $v_2[1]$ is Gaussian with zero mean and variance σ_{v2}^2, and $u_2[1] := f_2(y_2[1])$, $x[2] = x[1] + u_2[1]$. The control objective is minimizing $\mathbb{E}[qx[2]^2 + r_1u_1[0]^2 + r_2u_2[1]^2]$.

At a high level, this difference can be understood in terms of implicit communication. Radner's problem is a single-stage problem. Even if one controller sends some information, it is impossible for the other controller to receive the information at the next time step. Therefore, implicit communication between the controllers is impossible, and it is widely believed that if this is the case, then linear is optimal [109, 117, 92, 60, 83]. However, Witsenhausen's counterexample is a two-stage problem. If the first controller sends some information, the second controller can receive this information at the next time step. Therefore, implicit communication is possible, and nonlinear strategies which are good at this implicit communication can outperform linear strategies.

Let's revisit these problems using the binary deterministic models. Like in Section 4.3, we will give a perfect observation to the first controller and allow unbounded input power for the second controller. The goal of control is minimizing the state disturbance for a given input power constraint.

Binary deterministic model counterparts of Radner's problem and Witsenhausen's counterexample, shown in Figure 4.4b and 4.5b respectively, are formulated as follows.

Problem D (Binary Deterministic Model for Radner's Problem)**.** *For binary level index i, the deterministic system dynamics is given as follows:*

$$\begin{aligned}
x_i^1 &= x_i^0 \oplus u_{1,i}^0 \oplus u_{2,i}^0, \\
y_{1,i}^0 &= x_i^0, \\
y_{2,i}^0 &= x_i^0 \oplus v_i^0.
\end{aligned}$$

Here, x_i^0 are 0 for all $i \geq 2$ and Bernoulli $\frac{1}{2}$ on $\{0,1\}$ for all $i < 2$. v_i^0 are 0 for all $i \geq 1$ and Bernoulli $\frac{1}{2}$ on $\{0,1\}$ for all $i < 1$. $u_{1,i}^0$ and $u_{2,i}^0$ are functions of $y_{1,i}^0$ and $y_{2,i}^0$ respectively. The first controller has a power limit, $u_{1,i}^0 = 0$ for all $i \geq 1$.

The goal of the control is to minimize the final state distortion level d, i.e. minimizing d such that $x_i^1 = 0$ for all $i \geq d$.

Here, we can easily notice that $x_i^0, x_i^1, u_{1,i}^0, u_{2,i}^0, y_{1,i}^0, y_{2,i}^0, v_i^0$ correspond to $x[0]$, $x[1]$, $u_1[0]$, $u_2[0]$, $y_1[0]$, $y_2[0]$, $v_2[0]$ in the original Radner's problem.

Problem E (Binary Deterministic Model for Witsenhausen's Counterexample [37])**.** *For binary level index i, the deterministic system dynamics is given as follows:*

$$x_i^1 = x_i^0 \oplus u_{1,i}^0$$
$$x_i^2 = x_i^1 \oplus u_{2,i}^1$$
$$y_{1,i}^0 = x_i^0$$
$$y_{2,i}^1 = x_i^1 \oplus v_i^1$$

Here, x_i^0 are 0 for all $i \geq 2$ and Bernoulli $\frac{1}{2}$ on $\{0,1\}$ for all $i < 2$. v_i^1 are 0 for all $i \geq 1$ and Bernoulli $\frac{1}{2}$ on $\{0,1\}$ for all $i < 1$. $u_{1,i}^0$ and $u_{2,i}^1$ are functions of $y_{1,i}^0$ and $y_{2,i}^1$ respectively. The first controller has a power limit, $u_{1,i}^0 = 0$ for all $i \geq 1$.

The goal of the control is to minimize the final state distortion level d, i.e. minimizing d such that $x_i^2 = 0$ for all $i \geq d$.

Here, we can easily notice that $x_i^0, x_i^1, x_i^2, u_{1,i}^0, u_{2,i}^1, y_{1,i}^0, y_{2,i}^1, v_i^0$ correspond to $x[0]$, $x[1]$, $x[2]$, $u_1[0]$, $u_2[1]$, $y_1[0]$, $y_2[1]$, $v_2[1]$ in the original Witsenhausen's problem.

As we can see in Figure 4.4b and Figure 4.5b, essentially the same scheme that we discussed in Section 4.3 can be used in both deterministic problems to give the optimal cost. The first controller cancels the lower bits x_{20} and the second controller cancels the higher bits x_{10} at the next time step.

Proposition 4.5. *At time n, let the first controller's observation be $y_1[n] = \cdots y_{1,1}^n y_{1,0}^n . y_{1,-1}^n \cdots$ in binary expansion. Likewise, the second controller's observation is $y_2[n] = \cdots y_{2,1}^n y_{2,0}^n . y_{2,-1}^n \cdots$ in binary expansion. Then, the following control strategy achieves $d = -\infty$ (i.e. the final state is identically zero.) in both Problem D and E and is optimal in both problems.*

$$u_1[n] = y_{1,0}^n . y_{1,-1}^n \cdots$$
$$u_2[n] = \cdots y_{2,2}^n y_{2,1}^n 0.00 \cdots$$

Proof. Immediately follows from Figure 4.4b and Figure 4.5b. □

As we discussed in Section 4.3 the corresponding strategy in the reals is a nonlinear strategy. However, linear is optimal in Radner's problem. How can this be? Clearly, the real nonlinear strategy is not even approximately achieving the cost that the binary deterministic model promises. The binary deterministic model **fails to predict** the optimal control strategy and the optimal cost of the real Gaussian Radner's problem. The reason for this is the binary deterministic model ignores the carry-over in addition and subtraction which is actually happening in real Gaussian problems.

In fact, we can see the difference between $y_2[0]$ of Figure 4.4b and $y_2[1]$ of Figure 4.5b. The second bit of $y_2[0]$ of Figure 4.4b is $x_{20} \oplus v_{10}$ which causes carry-over in the reals, while the second bit of $y_2[1]$ of Figure 4.5b is v_{10} is just v_{10}. Therefore, the bitwise separation ignoring the carry-over results in an overly optimistic conclusion in binary deterministic models. A linear view of nonlinearity is too simplified in this case.

In fact, even in Witsenhausen's counterexample there is a small gap between the predicted cost and actual LQG cost, even though the deterministic model correctly predicts the approximately optimal strategy. As we can see in Proposition 4.5, in the deterministic model the final state is 0 as long as the first controller's input power is greater than the second controller's noise level. In the corresponding LQG problem, the final cost turns out to be only an exponentially decreasing function of the first controller's input power. However, the underlying reason for this gap is different from that in Radner's problem. This gap in Witsenhausen's counterexample comes from the tail of Gaussian random variables and the finite-dimensionality of the problem.[7] While all disturbances are bounded with probability 1 in deterministic models, in LQG problems Gaussian random variables can be arbitrary large with an exponentially decreasing probability. This results in a logarithmic gap between the costs in Witsenhausen's counterexample. However, unlike in Radner's problem this gap is only logarithmic and the insights that we gain from the deterministic models are still useful in the original LQG problems.

Therefore, we can rightfully say that deterministic models predict the essential behavior of Witsenhausen's counterexample, while failing for Radner's problem.

To clarify this point, we propose another simple deterministic model, the *ring model*, that takes into account of the carry-over effect. As shown in Figure 4.6a, there are 9 possible states, each time 1 is added the state rotates one step in counter clockwise, and each time 1 is subtracted the state rotates one step in clockwise. The distance between two states are measured by a minimum number of $+1$ or -1 that we have to add to move from one state to the other state. The norm of a state is defined as the distance from 0 to the state.

Let's apply this ring model to Witsenhausen's counterexample. We will consider the corresponding situation of the binary deterministic model in Figure 4.4b. $x[0]$ is uniformly random between all possible 9 states. At the fist controller, $y_1[0] = x[0]$ but $u_1[0] \in \{-1, 0, 1\}$. At the second controller, $y_2[1]$ is either $x[1] + 1$ or $x[1]$ or $x[1] - 1$ with probability $\frac{1}{3}$ and $u_2[1]$ can take arbitrary value. The goal of the control is minimizing the norm of the final state $x[2]$.

Figure 4.6b and 4.6c shows the optimal strategy for the first and second controller respectively, which is canceling lower and higher bits of uncertainty. As shown in Figure 4.6d, after the first controller's control $x[1]$ has only three possible states, $0, 3, 6$. Even after corruption by

[7] In infinite-dimensional problems, the laws of large numbers guarantee that Gaussian random variables behave typically and the probability that they can be arbitrary large asymptotically goes to zero. Therefore, we can drive the final cost to 0 with bounded first controller's input power, and the cost predicted by the deterministic model is actually achievable. To capture the finite-dimentionality of the problem, we have to use large deviation ideas. We refer to [38, 18] for further details.

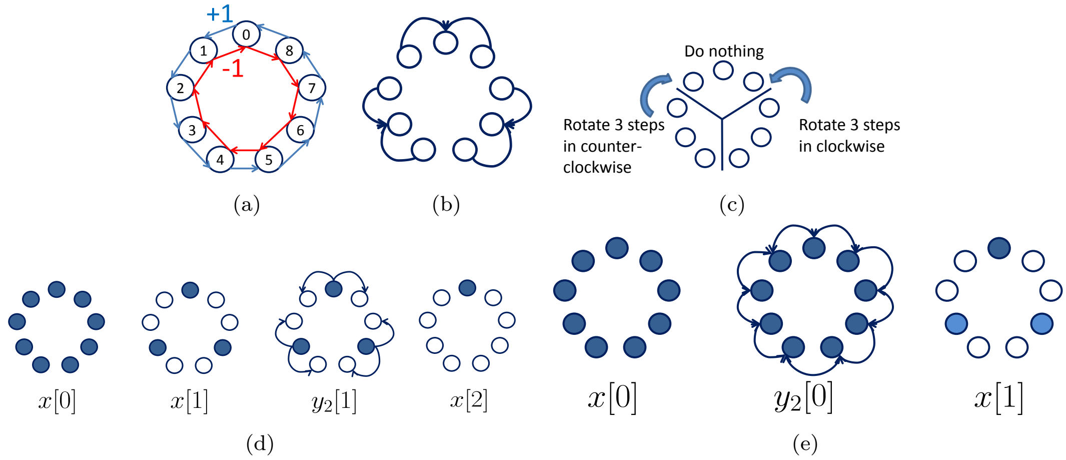


Figure 4.6: (a) Ring Model with 9 states. (b) A strategy for the first controller in ring model. (c) A strategy for the second controller in ring model. (d) Resulting state evolution for Witsenhausen's counterexample. (e) Resulting state evolution for Radner's problem.

the observation noise, $y_2[1]$ still has enough information to decode $x[1]$. Therefore, by the second controller's strategy in Figure 4.6c, the final state $x[2]$ can be forced to 0.

Then, let's apply the same strategy to Radner's problem. As we can see in Figure 4.6e, $x[0]$ is not quantized and the second controller cannot decode the initial state from its observation $y_2[0]$. The same strategy of Figure 4.6b and 4.6c gives a different result from Witsenhausen's counterexample. The final state $x[0]$ is 0 with probability $\frac{7}{9}$, 3 with probability $\frac{1}{9}$, and 6 with probability $\frac{1}{9}$. Thus, the average squared norm of $x[2]$ is 2. Let's consider a different strategy, $u_1[0] = 0$ and $u_2[0] = -y_2[0]$, which corresponds to a linear strategy in the Gaussian reals. Then, we can easily check that the final state $x[1]$ is equiprobably 1 or 0 or -1 and the average squared norm is $\frac{2}{3}$. Therefore, this linear strategy performs better than the nonlinear strategy that works for Witsenhausen's counterexample.

4.4 Proofs and Proof Ideas: High-Level Outline

The formal proof of the main result is separated into three parts. We will give upper and lower bounds on the optimal cost, and then compare them to show that they are within a constant ratio.

Figure 4.7 shows the proof idea flow for the upper bound[8] on the optimal cost. This is done by analyzing specific control strategies. First, it is easy to analyze linear strategies by simply tracking mean and variance. For nonlinear strategies, it can be tricky since mean and variance do not characterize non-Gaussian random variables. Therefore, in Section 4.5.2, we will introduce

[8]This corresponds to achievability arguments in information theory.

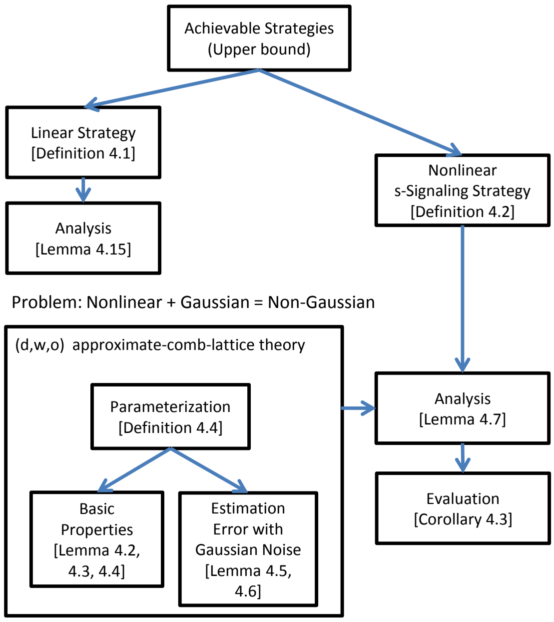


Figure 4.7: Flow diagram of the ideas for the upper bound on the control performance

a mini theory to analyze quantization-based strategies, which we call (d, w, o)-approximate-comb-lattice theory. Section 4.5.3 will actually analyze the nonlinear strategy performance based on this theory.

To show that we cannot do much better, we also have to find a lower bound[9] on the cost. Figure 4.8 shows the flow of ideas in the proof for the lower bound. The key idea is identifying the informational bottleneck of a problem and figuring out the information relaying between the controllers. In information theory, to figure out the informational bottleneck of the system, we partition the nodes and apply cutset bounds [6, 21]. However, here rather than simply partitioning the nodes, we expand the system in time and must divide the infinite-horizon problem into finite-horizon ones. The geometric slicing idea (Figure 4.15) is introduced for this.

Now, we have a finite-horizon problem. However, unlike infinite-horizon problems where the effect of transients can be amortized over infinitely many stationary states, the transient states are the essence of a finite-horizon problem and therefore the problem is non-stationary. To handle this issue, we divide the resulting finite-horizon problem into three sub time-intervals — childhood, youth and old age, so to speak. Figure 4.16 (or Figure 4.23) shows the division of time intervals. In “childhood”, we do not have enough information about the state, so we will call this interval information-limited. In “old age”, we do not have enough power to control the state too well, so we will call this interval power-limited. Between these two —in “youth”— something interesting is happening and we will call this interval a MIMO Witsenhaussen’s (or Radner’s for Figure 4.23)

[9]This corresponds to converse arguments in information theory.

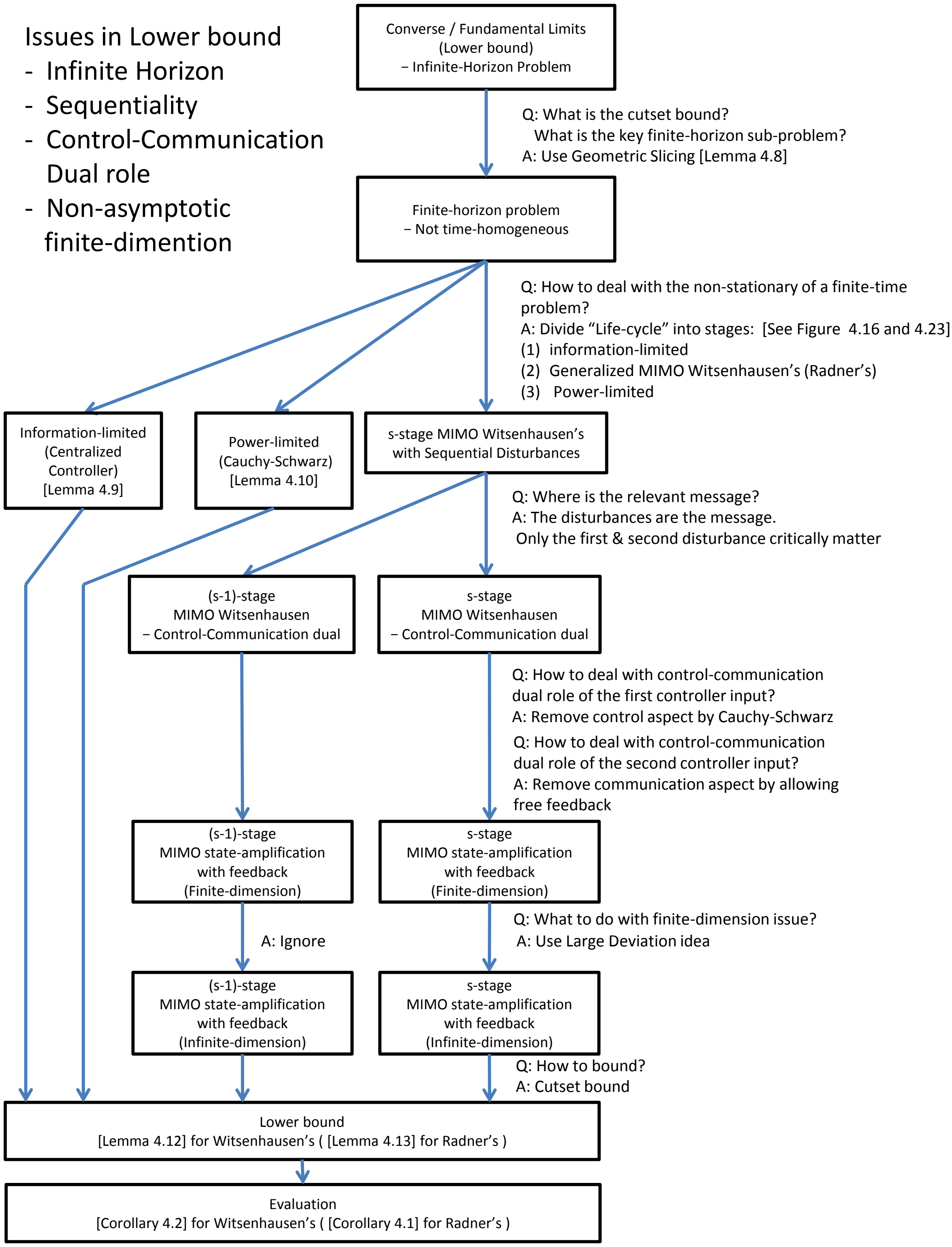


Figure 4.8: Flow diagram of the ideas for the lower bound on the control performance

interval.

In this interesting interval, the first controller is power-limited and the second controller is information-limited, which is essentially the same issue as in Witsenhausen's counterexample. In fact, we will relate this interval to an s-stage MIMO Witsenhausen's counterexample where a new disturbance is added at each time step. Then, the question becomes what are the critical disturbances among these? We will see that only first and second disturbances matter, and we can relax to simpler problems which are s-stage and $(s-1)$-stage MIMO Witsenhausen's with only one disturbance. However, still these problems are difficult due to the dual role of controller actions. The controller actions can be used to control the states, but at the same time they can be used to communicate some information to the other controller. This control-communication dual role of controller actions makes the problem hard.

To tackle this issue, we remove the control role from the first controller, and thereby the first controller will behave like a transmitter in communication problems. On the other hand, we remove the communication role from the second controller by allowing free feedback, and thereby the second controller will behave like a receiver in communication problems. In this way, we can reduce the problem to MIMO state-amplification with feedback, which generalizes the problem shown in [50]. However, the resulting problem is finite dimensional, and information-theoretic results for infinite-dimensional problems can possibly give loose bounds [37]. In fact, we have to adapt large deviation ideas to the s-stage MIMO state-amplification problem for this reason.[10] Now, we can apply simple information-theoretic cutset bounds to the final communication problems and derive lower bounds.

Before we discuss the proof details, we first convert the weighted long-term average cost optimization problem to an optimization problem with average power constraints as we did in Section 4.3. The original control objective is minimizing the weighted cost of the state disturbance and the controller input powers. However, it will be useful to consider minimizing the state given an average bound on the input powers. Formally, the problem is written as follows.

Problem F (Decentralized LQG problem with average power constraints). *Consider the same dynamics as Problem B. But, now the control objective is minimizing the state disturbance $D(P_1, P_2)$ for given input power constraints $P_1, P_2 \in \mathbb{R}^+$. We will say the power-disturbance tradeoff,* $D(P_1, P_2)$

[10] This is the same issue and idea that we discussed in Section 4.3.1 for Witsenhausen's counterexample and the issue addressed in [37]

is achievable if and only if there exist causal control strategies $u_1[n], u_2[n]$ *such that*

$$\limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mathbb{E}[x^2[n]] \le D(P_1, P_2),$$
$$\limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mathbb{E}[u_1^2[n]] \le P_1,$$
$$\limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mathbb{E}[u_2^2[n]] \le P_2.$$

Lemma 4.14 will relate the weighted-cost problem, Problem B, and the power-constraints problem, Problem F, telling us that if we can approximately solve the latter we can also approximately solve the former. To characterize $D(P_1, P_2)$ approximately, we will come up with lower and upper bounds on $D(P_1, P_2)$. Since we are only aiming for an approximate solution, in the discussion for intuitions and interpretations we will focus on the scaling and ignore the constants.

The following Cauchy-Schwarz style inequalities will be helpful to get bounds.

Lemma 4.1. *For arbitrarily correlated random variables* $X_1, \cdots, X_n$*, the following inequality holds:*

$$(\sqrt{\mathbb{E}[X_1^2]} - \sqrt{\mathbb{E}[X_2^2]} \cdots - \sqrt{\mathbb{E}[X_n^2]})_+^2 \le \mathbb{E}[(X_1 + \cdots + X_n)^2] \le (\sqrt{\mathbb{E}[X_1^2]} + \cdots + \sqrt{\mathbb{E}[X_n^2]})^2$$
$$\le n(\mathbb{E}[X_1^2] + \cdots + \mathbb{E}[X_n^2])$$

Proof.

$$\begin{aligned}
&\mathbb{E}[(X_1 + \cdots + X_n)^2] \\
&= \mathbb{E}[X_1^2] + \cdots + \mathbb{E}[X_n^2] + 2\mathbb{E}[X_1X_2] + \cdots + 2\mathbb{E}[X_{n-1}X_n] \\
&\le \mathbb{E}[X_1^2] + \cdots + \mathbb{E}[X_n^2] + 2\sqrt{\mathbb{E}[X_1^2]\mathbb{E}[X_1^2]} + \cdots + 2\sqrt{\mathbb{E}[X_1^2]\mathbb{E}[X_n^2]} \\
&= (\sqrt{\mathbb{E}[X_1^2]} + \cdots + \sqrt{\mathbb{E}[X_n^2]})^2 \\
&\le n(\mathbb{E}[X_1^2] + \cdots + \mathbb{E}[X_n^2])
\end{aligned}$$

where all inequalities follow from Cauchy-Schwarz.

$$\begin{aligned}
&\mathbb{E}[(X_1 + \cdots + X_n)^2] \\
&= \mathbb{E}[X_1^2] + 2\mathbb{E}[X_1(X_2 + \cdots + X_n)] + \mathbb{E}[(X_2 + \cdots + X_n)^2] \\
&\ge \mathbb{E}[X_1^2] - 2\sqrt{\mathbb{E}[X_1^2]\mathbb{E}[(X_2 + \cdots + X_n)^2]} + \mathbb{E}[(X_2 + \cdots + X_n)^2] \\
&= (\sqrt{\mathbb{E}[X_1^2]} - \sqrt{\mathbb{E}[(X_2 + \cdots + X_n)^2]})^2 \\
&\ge (\sqrt{\mathbb{E}[X_1^2]} - \sqrt{\mathbb{E}[X_2^2]} \cdots - \sqrt{\mathbb{E}[X_n^2]})_+^2
\end{aligned}$$

where all inequalities follow from Cauchy-Schwarz. □

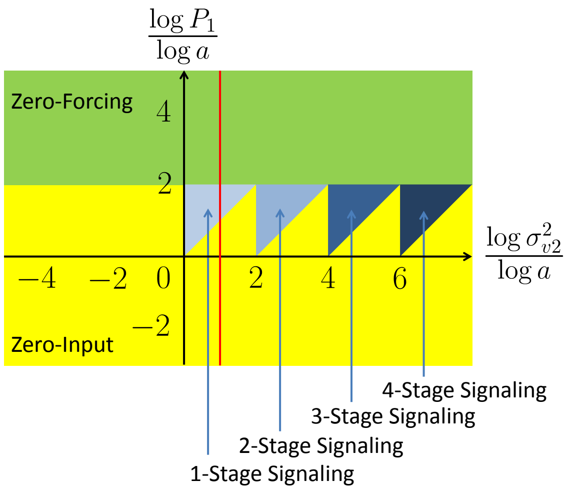


Figure 4.9: Approximately optimal strategies for given P_1 and σ_{v2}^2 when $\sigma_{v1}^2 = 0$ and $P_2 = \infty$.

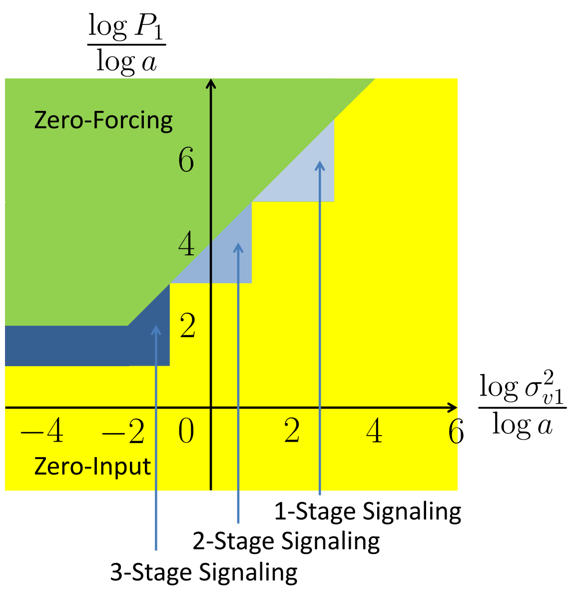


Figure 4.10: Approximately optimal strategies for given P_1 and σ_{v1}^2 when $\sigma_{v2}^2 = a^5$ and $P_2 = \infty$.

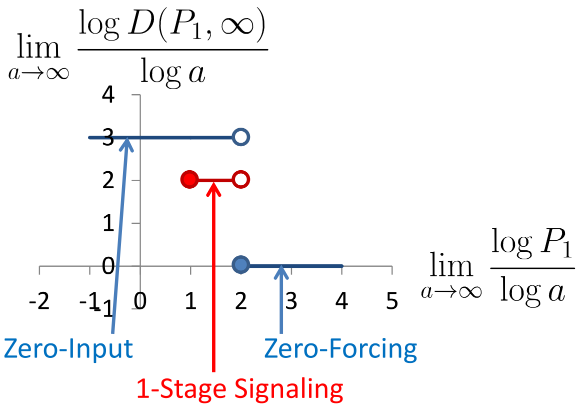


Figure 4.11: The minimum state disturbance $D(P_1, P_2)$ when $\sigma_{v1}^2 = 0$, $\sigma_{v2} = a$ and $P_2 = \infty$ as a function of P_1. The red line indicates the cost achievable by the 1-stage signaling strategy. The blue line indicates the cost achievable by linear strategies. As we can see this performance plot corresponds to that of the red line in Figure 4.9.

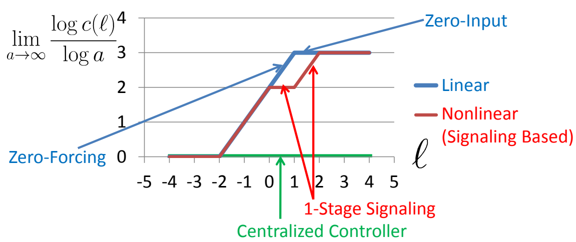


Figure 4.12: The optimal weighted average cost for $\sigma_{v1}^2 = 0$, $\sigma_{v2}^2 = a$, $q = 1$, $r_1 = a^l$, $r_2 = 0$. The red line indicates the optimal cost among all possible strategies. The blue line indicates the optimal cost among only linear strategies. The green line indicates the cost of the centralized controller which has both observations and can control both inputs. As l varies, the optimal strategy traverses the red line of Figure 4.9.

4.5 Proofs and Proof Ideas: Upper bound on the optimal cost

To come up with an upper bound on $D(P_1, P_2)$, we should propose appropriate achievable control strategies for approximate optimality and analyze their performances.

As we discussed in Section 4.3.1, a 1-stage signaling strategy ($L_{sig,1}$) for the infinite-horizon problem (shown in Figure 4.2) and the nonlinear strategy for Witsenhausen's counterexample (shown in Figure 4.5) are essentially equivalent. The first controller implicitly communicates its observation to the second controller by forcing the lower state bits to be zero. This point can be visually understood in Figure 4.1 by noticing that Witsenhausen's counterexample is indeed a sub-block of the infinite-horizon problem.

However, there is a significant difference between these two problems — the time-horizon. Witsenhausen's counterexample terminates after 2-time steps, while the system keeps running in infinite-horizon problems. Therefore, more issues arise when we are designing controllers for infinite-horizon problems.

First, since the system keeps running in infinite horizon problems, the implicit communication also has to keep happening. In Figure 4.1, $C_1[1]$ communicates to $C_2[2]$, $C_1[2]$ communicates to $C_2[3]$, and so on. In other words, an infinite-horizon problem can be thought as a series of Witsenhausen counterexamples. Because of this interlocking of Witsenhausen's blocks, the effect of one problem can propagate to subsequent ones. To handle this interference between interlocked problems, we introduced the $R_{a^s d}(\sum_{1 \leq i \leq s} a^{i-1} u_2[n-i])$ terms in the s-stage signaling policy in Definition 4.2.

The second difference is that since we have a longer time horizon, $C_1[0]$ does not have to communicate to $C_2[1]$ of the next time step. It can also communicate with longer delay to $C_2[2]$, $C_2[3]$, $\cdots$. In general, $C_1[0]$ can communicate to $C_2[s]$ as we can see in Figure 4.1. In fact, the s-stage signaling strategy of Definition 4.2, $L_{sig,s}$, enables $C_1[0]$ to communicate with $C_2[s]$, and the infinite-horizon problem is decomposed into a series of interlocked 's-stage MIMO Witsenhausen's counterexamples'.

Let's take a careful look at these signaling strategies, and understand which strategy has to be used for which parameters of Problem B. For simplicity, we first consider the extreme case when the first controller has a perfect observation and the second controller has no power constraint just like Section 4.3. In other words, $\sigma_{v1}^2 = 0$ and $P_2 = \infty$. Here, we will be making references to the binary deterministic perspective on the problem.

When $\sigma_{v1}^2 = 0$ and $P_2 = \infty$

Figure 4.9 summarizes which strategy has to be used for a given σ_2^2 and P_1. First, we can notice that if the first controller has enough power then it does not really need any help from the second controller. At each time step the disturbance $w[n]$ is added, it is observable at the next time step $n+1$ by the first controller when its power is amplified by a^2. Therefore, if $P_1 \geq a^2$ the first controller can remove the disturbance by itself by choosing $u_1[n] = -ay_1[n]$. We will call this a zero-forcing strategy from the first controller's point of view. On the other hand, at each time a new state disturbance $w[n]$ with variance 1 is added. Therefore, when $P_1 \leq 1$ most of the first controller's input will be masked by the additional disturbance $w[n]$. Therefore, in this case $u_1[n] = 0$ is approximately optimal, and we will call this a zero-input strategy from the first controller's point of view.

Therefore, the question is "what should the first controller do when P_1 is between these two extreme values?" As we discussed before, the first controller can implicitly communicate its perfect observation to the second controller by canceling the bits which are not observable by the second controller. This idea can be implemented when the bits below the second controller's noise level are observed by the first controller at previous time steps. For example, in Figure 4.2, x_{-2}^1 of time step 2, the bit below the noise level of the second controller, is observed by the first controller at time step 1, one time step before.

Then, what is the condition for the first controller to observe the disturbance one time step before in the original LQG problems? We can notice that at each time the disturbance $w[n]$ is amplified by a and its variance becomes a^2 after one time step. Therefore, when $1 \leq \sigma_{v2}^2 \leq a^2$ the bits below the second controller's noise level are observed by the first controller at 1 time step before.

What is the minimum power required for the first controller to cancel all the bits below the second controller's noise level σ_{v2}^2? As we can guess[11], the answer is σ_{v2}^2. In sum, for 1-stage signaling to be actually useful, the parameters of the LQG problems have to be $1 \leq \sigma_{v2}^2 \leq a^2$ and $\sigma_{v2}^2 \leq P_1 \leq a^2$. When $P_1 \geq a^2$, zero-forcing is approximately optimal, and when $0 \leq P_1 \leq \sigma_{v2}^2$, zero-input is approximately optimal.

In general, when $a^{2(s-1)} \leq \sigma_{v2}^2 \leq a^{2s}$ for some $s \in \mathbb{N}$, the bits below the second controller's noise level can be "previewed" by the first controller at s time steps before, and the first controller's power has to be larger than $\frac{\sigma_{v2}^2}{a^{2(s-1)}}$ to actually cancel those bits. Therefore, in this case when $P_1 \geq a^2$, zero-forcing is approximately optimal, when $\frac{\sigma_{v2}^2}{a^{2(s-1)}} \leq P_1 \leq a^2$, s-stage signaling is approximately optimal, and when $0 \leq P_1 \leq \frac{\sigma_{v2}^2}{a^{2(s-1)}}$, zero-input is approximately optimal.

On the other hand, when $\sigma_{v2}^2 \leq 1$, it corresponds to dividing the infinite-horizon problem into a series of Radner's problems.[12] The first controller will observe the bits below the second

[11] We can also conjecture this from the deterministic model in Figure 4.2. In Figure 4.2, the first controller's input power level and the second controller's noise level is the same.

[12] In Section 4.7, we will name this case as the weakly-degraded-observation case, while the remaining case is named

controller's noise level without any delay, so it gets no preview. As we discussed in Section 4.3.1, we cannot expect a significant gain from nonlinear strategies when two controllers are acting simultaneously on essentially the same quality observations. Therefore, in this case, a linear strategy is enough to achieve constant-ratio optimality. We will revisit this point when we are discussing lower bounds in Section 4.6.

When $\sigma_{v1}^2 > 0$

So far, we limited ourselves to $\sigma_{v1}^2 = 0$ and $P_2 = \infty$. Let's first consider the case when $\sigma_{v1}^2 > 0$.

If we take a careful look at the previous case of $\sigma_{v1}^2 = 0$, the bits that the first controller actually uses are those between the power level 1 and a^{-2}. The bits below a^{-2} are useless since at the next time step, they will be masked by the new disturbance. Therefore, as long as $\sigma_{v1}^2 \leq a^{-2}$, the first controller can observe all its useful bits and the previous argument does not change.

Then, what is happening in the case when $\sigma_{v1}^2 \geq a^{-2}$? First, let's ask what is the minimum power P_1 for the first controller to zero-force the state. The disturbance is amplified by a^2 at each time step, and the bits below σ_{v1}^2 are not observable by the first controller. Therefore, by the time the first controller observes the effect of the disturbance, the state's variance becomes $a^2\sigma_{v1}^2$. To actually cancel it at the next time step, the first controller's power has to be greater than $a^4\sigma_{v1}^2$, i.e. $P_1 \geq a^4\sigma_{v1}^2$.

When the input power is smaller than $a^4\sigma_{v1}^2$, it has to use signaling strategies. So when can we use the s-stage signaling strategy? To use the s-stage signaling strategy, the first controller has to observe the bits below the second controller's noise level at least s time steps before. Therefore, σ_{v1}^2 has to be less than $\frac{\sigma_{v2}^2}{a^{2s}}$. Since a longer stage signaling requires smaller power, we will use an s-stage signaling strategy when $\frac{\sigma_{v2}^2}{a^{2(s+1)}} < \sigma_{v1}^2 \leq \frac{\sigma_{v2}^2}{a^{2s}}$. Then, what is the minimum power to use s-stage signaling strategy? Since the first controller has to cancel the bits below $\frac{\sigma_{v2}^2}{a^{2s}}$ at the next time step, P_1 has to be greater than $\frac{a^2\sigma_{v2}^2}{a^{2s}}$. When P_1 is less than this, the first controller uses the zero-input strategy.

Summarizing the conclusions so far, let $s = \lceil \frac{\ln \sigma_{v2}^2 - \ln(\max(1, a^2\sigma_{v1}^2))}{2\ln a} \rceil$ so that

$$a^{2(s-1)} \max(1, a^2\sigma_{v1}^2) < \sigma_{v2}^2 \leq a^{2s} \max(1, a^2\sigma_{v1}^2).$$

Then (i) When $P_1 \geq \max(a^2, a^4\sigma_{v1}^2)$, the zero-forcing strategy
(ii) When $\frac{\sigma_{v2}^2}{a^{2(s-1)}} \leq P_1 \leq \max(a^2, a^4\sigma_{v1}^2)$, the s-stage signaling strategy
(iii) When $P_1 \leq \frac{\sigma_{v2}^2}{a^{2(s-1)}}$, the zero-input strategy
are approximately optimal respectively.

---

the strongly-degraded-observation case.

When $P_2 < \infty$

Let's consider when the second controller also has a power constraint P_2. When the first controller is zero-forcing the state, the second controller does not have to control and the power constraint P_2 does not change the result. When the first controller is either applying signaling or zero input strategy, the second controller has to stabilize the system. By the definition, $D(P_1, \infty)$ is the smallest state disturbance we can expect. Therefore, P_2 has to be essentially greater than $a^2 D(P_1, \infty)$ to cancel the state at the next time step and stabilize the system. In fact, this turns out to be sufficient, too.

4.5.1 Generalized d.o.f. Performance

Now, we have approximately optimal strategies. In this section, we will see how the performance scales as the problem parameters vary. More precisely, we will increase the various problem parameters in different scales, and see how the control cost scales as a function of the problem parameters. In spirit, this measure of the performance corresponds to the generalized d.o.f. in wireless communication [29, 6] where the SNRs of different antennas are allowed to scale differently. The more fundamental connection with wireless communication theory will be discussed in Section 4.8.

Figure 4.11 shows how the minimum state disturbance of the proposed strategies scales as a goes to infinity. Precisely, in Problem F we fix $a = a$, $\sigma_{v1}^2 = 0$, $\sigma_{v2}^2 = a$, $P_2 = \infty$, and explore how $D(P_1, P_2)$ scales in a when P_1 scales differently in a. From the problem parameters, we can easily see this cost plot corresponds to the cost of the red line ($\sigma_{v2}^2 = a$) in Figure 4.9. As we discussed before, between zero-forcing and zero-input linear strategies, the nonlinear 1-stage signaling strategy performs better.

So far the discussion is from the power-disturbance point of view. However, the original weighted cost problem is essentially the same since the optimal strategy will have some corresponding control input powers. Let's consider the system equation (4.2) with $a = a, \sigma_{v1}^2 = 0, \sigma_{v2}^2 = a$ and the average cost (4.1)

$$c(l) = \inf_{u_1, u_2} \limsup_{N \to \infty} \frac{1}{N} \sum_{0 \leq n < N} \mathbb{E}[x^2[n] + a^l u_1^2[n]]$$

Figure 4.12 shows how the average cost scales as a goes to infinity for different values of l. As we change l, the optimal solution follows the red line ($\sigma_{v2}^2 = a$) in Figure 4.9.

(i) When l is small ($l \leq 0$), the input cost of the first controller is inexpensive and the zero-forcing strategy is optimal up to scaling.

(ii) When l is large ($l \geq 2$), the input cost is expensive and the zero-input strategy is approximately optimal.

(iii) Between these two extremes ($0 \leq 1 \leq 2$), we need a nonlinear 1-stage signaling strategy and it is approximately optimal.

As we can see in Figure 4.12, the average cost of linear and optimal nonlinear strategy scales differently in a. Therefore, the performance ratio between these two diverges to infinity, which was formally stated in Proposition 4.4. Moreover, in Figure 4.12 we can also see a naive lower bound on the cost (derived by allowing a centralized controller) is too loose to give constant ratio optimality. Thus, we have to improve both the upper and lower bounds to prove constant-ratio optimality.

It is worth mentioning that figuring out this generalized d.o.f. cost is not enough to guarantee constant-ratio optimality, since the logarithmic scaling (caused by the tail of the Gaussian random variables) in a does not appear in the generalized d.o.f. cost. For example, the first term shown in the lower bound given by (c) of Corollary 4.2 cannot be captured in the generalized d.o.f. cost.

4.5.2 (d,w,o) Approximate-Comb-Lattice Theory

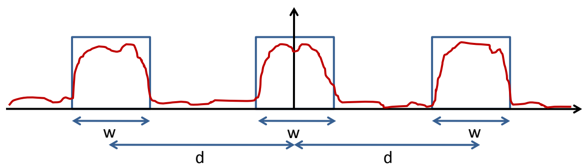


Figure 4.13: Pictorial description of Definition 4.4

So far, we understand the approximately optimal strategies and intuitively why they have to be used for given problem parameters. Now, we have to formally analyze their performances. Unlike linear strategies, nonlinear strategies make the random variables (the state, observations and inputs) non-Gaussian. Thus, the mean and variance is not enough to describe the distribution of random variables, and the exact description of the distribution requires a potentially infinite number of parameters. Therefore, we have to come up with an approximate description involving only a finite number of parameters. To this end, we propose the following definition which will turn out to be useful in analyzing quantization-based signaling strategies.

Definition 4.4. *Let X be a random variable, d be nonzero, and w, o be nonnegative reals with $|d| > w$. We say $X \leq_{df} (d, w, o)$ if*

$$\mathbb{P}\{X \notin \bigcup_{i \in \mathbb{Z}} [i \cdot d - \frac{w}{2}, i \cdot d + \frac{w}{2}]\} \leq o.$$

Figure 4.13 pictorially shows this definition. When a random variable stays in one of the boxes with width w, the event will be considered typical. When a random variable falls outside the boxes, the event will be considered atypical and measured by outage probability o. Notice that after quantization, the probability mass will concentrate in a sequence of boxes. When $d = \infty$, we have only one box centered at 0.

Let's study properties of this definition. The first lemma tells what happens when we add two random variables.

Lemma 4.2. *Let X_1, X_2 and X_3 be arbitrary correlated random variables. If $X_1 \leq_{df} (d_1, w_1, o_1)$, $X_2 \leq_{df} (\infty, w_2, o_2)$ and $X_3 \leq_{df} (d_1, w_3, o_3)$ then*

$$X_1 + X_2 \leq_{df} (d_1, w_1 + w_2, o_1 + o_2)$$
$$X_1 + X_3 \leq_{df} (d_1, w_1 + w_3, o_1 + o_3)$$

Proof.

$$\begin{aligned}
&\mathbb{P}\{X_1 + X_2 \notin \bigcup_{i\in\mathbb{Z}} [i \cdot d_1 - \frac{w_1 + w_2}{2}, i \cdot d_1 + \frac{w_1 + w_2}{2}]\} \\
&\leq \mathbb{P}\{X_1 + X_2 \notin \bigcup_{i\in\mathbb{Z}} [i \cdot d_1 - \frac{w_1 + w_2}{2}, i \cdot d_1 + \frac{w_1 + w_2}{2}], X_2 \in [-\frac{w_2}{2}, \frac{w_2}{2}]\} + \mathbb{P}\{X_2 \notin [-\frac{w_2}{2}, \frac{w_2}{2}]\} \\
&\leq \mathbb{P}\{X_1 \notin \bigcup_{i\in\mathbb{Z}} [i \cdot d_1 - \frac{w_1}{2}, i \cdot d_1 + \frac{w_1}{2}]\} + o_2 \\
&\leq o_1 + o_2
\end{aligned}$$

The second part follows similarly since when we add two points from the lattice points spaced by d, the resulting point is also in that lattice. □

The second lemma tells what happens when we multiply a random variable by a constant.

Lemma 4.3. *Let $X \leq_{df} (d, w, o)$ and $k > 0$. Then,*

$$kX \leq_{df} (kd, kw, o).$$

Proof.

$$\begin{aligned}
&\mathbb{P}\{kX \notin \bigcup_{i\in\mathbb{Z}} [i \cdot kd - \frac{kw}{2}, i \cdot kd + \frac{kw}{2}]\} \\
&= \mathbb{P}\{X \notin \bigcup_{i\in\mathbb{Z}} [i \cdot d - \frac{w}{2}, i \cdot d + \frac{w}{2}]\} \\
&\leq o.
\end{aligned}$$

□

The next lemma captures the fact that the variance of a remainder is only smaller than the original random variable.

Lemma 4.4. *For all random variable X and nonzero d, we have*

$$\mathbb{E}[R_d(X)^2] = \mathbb{E}[(X - Q_d(X))^2] \le \mathbb{E}[X^2].$$

Proof. For a real x, let $x = q \cdot d + r$ for $q \in \mathbb{Z}$ and $r \in [-\frac{d}{2}, \frac{d}{2})$. Then,

$$\begin{aligned} x^2 &= (q \cdot d + r)^2 = q^2 d^2 + 2qdr + r^2 \\ &= |qd|(|qd| + 2sgn(qd) \cdot r) + r^2 \end{aligned}$$

When $q = 0$, $x^2 = r^2$.
When $q \ne 0$, since $q \in \mathbb{Z}$ we have $x^2 \ge |qd|(|d| - 2|r|) + r^2 \ge r^2$.

Therefore,

$$\mathbb{E}[X^2] \ge \mathbb{E}[R_d(X)^2].$$

□

Since all underlying random variables of interest are Gaussian, we will relate Gaussian distributions with our parameterization.

Lemma 4.5. *Let $Q(x) := \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp(-\frac{u^2}{2}) du$. Then, $Q(x) \sim \frac{1}{\sqrt{2\pi}x} \exp(-\frac{x^2}{2})$. More precisely, for $\forall x > 0$*

$$\frac{1}{\sqrt{2\pi}} \left(\frac{1}{x} - \frac{1}{x^3} \right) \exp\left(-\frac{x^2}{2} \right) \le Q(x) \le \frac{1}{\sqrt{2\pi}x} \exp\left(-\frac{x^2}{2} \right).$$

Moreover, when X is Gaussian with zero mean and variance smaller than σ^2, for all $w \ge 0$

$$X \le_{df} (\infty, w, 2 \cdot Q(\frac{w}{2\sigma})).$$

Proof. For the first part, see [30]. The second part directly follows from the definition. □

The next lemma bounds the MMSE error of a quantized random variable when it is corrupted by Gaussian observation noise.

Lemma 4.6. *Let X and V be independent random variables where $X \le_{df} (d, w, o)$ with $|d| > w$ and V is a Gaussian random variable with zero-mean and variance σ^2. Then,*

$$\begin{aligned} &\mathbb{E}[(X - Q_d(X+V))^2] \\ &\le \mathbb{E}[(X - Q_d(X))^2] + \sum_{1 \le i \le \infty} (i|d| + \frac{w}{2})^2 \cdot 2Q(\frac{(2i-1)|d| - w}{2\sigma}) \\ &+ o \cdot \left((d + \frac{d}{2})^2 + \sum_{2 \le i \le \infty} (i \cdot d + \frac{d}{2})^2 \cdot 2Q_d(\frac{(i-1)|d|}{\sigma}) \right). \end{aligned}$$

Proof. For convenience, let $d > 0$. $d < 0$ can be proved by replacing d with $|d|$. Denote $\mathcal{T}_{d,w} = \bigcup_{i \in \mathbb{Z}} [i \cdot d - \frac{w}{2}, i \cdot d + \frac{w}{2}]$.

$$\begin{aligned}\mathbb{E}[(X - Q_d(X+V))^2] &= \mathbb{E}[(X - Q_d(X+V))^2 | X \in \mathcal{T}_{d,w}] \mathbb{P}\{X \in \mathcal{T}_{d,w}\} \\ &+ \mathbb{E}[(X - Q_d(X+V))^2 | X \in \mathcal{T}_{d,w}^c] \mathbb{P}\{X \in \mathcal{T}_{d,w}^c\} \\ &\leq \mathbb{E}[(X - Q_d(X+V))^2 | X \in \mathcal{T}_{d,w}] \\ &+ \mathbb{E}[(X - Q_d(X+V))^2 | X \in \mathcal{T}_{d,w}^c] \cdot o\end{aligned}$$

Notice that when $X \in \mathcal{T}_{d,w}$ and $|V| < \frac{d-w}{2}$, $Q_d(X) = Q_d(X+V)$. When $X \in \mathcal{T}_{d,w}$ and $|V| < d + \frac{d-w}{2}$, $Q_d(X) = Q_d(X+V) \pm d$ and so on. Therefore,

$$\begin{aligned}&\mathbb{E}[(X - Q_d(X+V))^2 | X \in \mathcal{T}_{d,w}] = \mathbb{E}[(Q_d(X) + R_d(X) - Q_d(X+V))^2 | X \in \mathcal{T}_{d,w}] \\ &\leq \mathbb{E}[(X - Q_d(X))^2] + (d + \frac{w}{2})^2 \cdot 2Q(\frac{d-w}{2\sigma}) + (2d + \frac{w}{2})^2 \cdot 2Q(\frac{3d-w}{2\sigma}) + \cdots\end{aligned}$$

Moreover, for all x when $|V| < d$, $Q_d(x) - Q_d(x+V) = -d, 0, d$. When $|V| < 2d$, $Q_d(x) - Q_d(x+V) = -2d, -d, 0, d, 2d$ and so on. Therefore, since $|R_d(\cdot)| \leq \frac{d}{2}$,

$$\begin{aligned}&\mathbb{E}[(X - Q_d(X+V))^2 | X \in \mathcal{T}_{d,w}^c] = \mathbb{E}[(Q_d(X) - Q_d(X+V) + R_d(X))^2 | X \in \mathcal{T}_{d,w}^c] \\ &\leq (d + \frac{d}{2})^2 + (2d + \frac{d}{2})^2 \cdot 2Q_d(\frac{d}{\sigma}) + (3d + \frac{d}{2})^2 \cdot 2Q_d(\frac{2d}{\sigma}) + \cdots\end{aligned}$$

Therefore,

$$\begin{aligned}&\mathbb{E}[(X - Q_d(X+V))^2] \\ &\leq \mathbb{E}[(X - Q_d(X))^2] + (d + \frac{w}{2})^2 \cdot 2Q(\frac{d-w}{2\sigma}) + (2d + \frac{w}{2})^2 \cdot 2Q(\frac{3d-w}{2\sigma}) + \cdots \\ &+ o \cdot ((d + \frac{d}{2})^2 + (2d + \frac{d}{2})^2 \cdot 2Q_d(\frac{d}{\sigma}) + (3d + \frac{d}{2})^2 \cdot 2Q_d(\frac{2d}{\sigma}) + \cdots).\end{aligned}$$

□

4.5.3 Analysis of Signaling Strategies

Now, we are ready to analyze the performance of the s-stage signaling strategy. In the s-stage signaling strategy, the first controller imposes a lattice structure on $x[n]$, but the second controller's action can possibly break this lattice structure. However, the second controller knows all its past control inputs, so it can exploit the imposed lattice structure by compensating for its past control inputs. More precisely, we will see that $x[n] - R_{a^s d}(\sum_{1 \leq i \leq s} a^{i-1} u_2[n-i])$ —with the compensation term, $R_{a^s d}(\sum_{1 \leq i \leq s} a^{i-1} u_2[n-i])$— has a lattice structure, and the second controller will observe this quantized state with an observation noise $v_2[n]$. In spirit, the idea and analysis in this section is similar to that in [78].

Before we state the lemma, we introduce a definition to compare multiple numbers. For $a_1, \ldots a_n, b_1, \ldots, b_n \in \mathbb{R}$, we say $(a_1, \ldots, a_n) \leq (b_1, \ldots, b_n)$ if and only if $a_1 \leq b_1$, …, $a_n \leq b_n$.

Lemma 4.7. *For a given $s \in \mathbb{N}$, let $S_{U,1}$ be the set of (d, w_1) such that*

$$d > 0, w_1 > 0,$$
$$|a|^s d - (|a|^{s-1} d \frac{|a|}{|a|-1} + w_1) > 0.$$

The bound $D_{U,1}(d, w_1)$ is defined as

$$\begin{aligned}
D_{U,1}(d, w_1) &:= 2a^{2s}(2(\frac{d}{2})^2(\frac{1}{1-\frac{1}{|a|}})^2 + 2(\frac{1}{1-\frac{1}{a^2}}) + 2a^2\sigma_{v1}^2) \\
&+ \sum_{1 \le i \le \infty} 4a^2(i|a|^s d + \frac{|a|^{s-1} d \frac{|a|}{|a|-1} + w_1}{2})^2 Q(\frac{(2i-1)|a|^s d - (|a|^{s-1} d \frac{|a|}{|a|-1} + w_1)}{2\sigma_{v2}}) \\
&+ 8a^2 Q(\frac{w_1}{2\sqrt{a^{2(s-1)}\frac{a^2}{a^2-1} + a^{2s}\sigma_{v1}^2}}) \sum_{1 \le i \le \infty} (i|a|^s d + \frac{|a|^s d}{2})^2 Q(\frac{(i-1)|a|^s d}{\sigma_{v2}}) \\
&+ 2(a^2(\frac{d}{2})^2) + 1.
\end{aligned} \tag{4.5}$$

Let $|a| > 1$. Then, for all s and $(d, w_1) \in S_{U,1}$, the s-stage signaling strategy of Definition 4.2 can achieve the following Power-Disturbance tradeoff of Problem F.

$$(D(P_1, P_2), P_1, P_2) \le (D_{U,1}(d, w_1), \frac{a^2 d^2}{4}, 8a^2 D_{U,1}(d, w_1) + \frac{7}{2} a^{2(s+1)} d^2 + 4a^2 \sigma_{v2}^2)$$

Proof. For notational simplicity, we only consider $a > 1$. The proof for $a < -1$ can be obtained by replacing a with $|a|$.

By the definition of s-stage signaling strategies,

$$\begin{aligned}
u_1[n] &= -aR_d(y_1[n]) \\
u_2[n] &= -a(Q_{a^s d}(y_2[n] - R_{a^s d}(\sum_{1 \le i \le s} a^{i-1} u_2[n-i])) + R_{a^s d}(\sum_{1 \le i \le s} a^{i-1} u_2[n-i]))
\end{aligned}$$

Therefore, for all n we have

$$\begin{aligned}
x[n+1] &= ax[n] + u_1[n] + u_2[n] + w[n] \\
&= ax[n] - a(Q_{a^s d}(y_2[n] - R_{a^s d}(\sum_{1 \le i \le s} a^{i-1} u_2[n-i])) + R_{a^s d}(\sum_{1 \le i \le s} a^{i-1} u_2[n-i])) + u_1[n] + w[n] \\
&= a(\underbrace{x[n] - R_{a^s d}(\sum_{1 \le i \le s} a^{i-1} u_2[n-i])}_{:=X[n]} - Q_{a^s d}(\underbrace{y_2[n] - R_{a^s d}(\sum_{1 \le i \le s} a^{i-1} u_2[n-i])}_{:=Y_2[n]})) + u_1[n] + w[n]
\end{aligned} \tag{4.6}$$

First, we will prove that for all $n \ge s$, $X[n]$ has a lattice structure. Then, $Y_2[n]$ is $X[n] + v_2[n]$, so we can use Lemma 4.6 to analyze the estimation error of quantized random variables.

For $n \geq s$, we have

$$
\begin{aligned}
&X[n] \\
&= x[n] - R_{a^s d}(\sum_{1\leq i\leq s} a^{i-1}u_2[n-i]) \\
&= a^s x[n-s] + \sum_{1\leq i\leq s} a^{i-1}u_1[n-i] + \sum_{1\leq i\leq s} a^{i-1}u_2[n-i] + \sum_{1\leq i\leq s} a^{i-1}w[n-i] - R_{a^s d}(\sum_{1\leq i\leq s} a^{i-1}u_2[n-i]) \\
&= (a^s x[n-s] + a^{s-1}u_1[n-s]) + \sum_{1\leq i\leq s} a^{i-1}u_2[n-i] - R_{a^s d}(\sum_{1\leq i\leq s} a^{i-1}u_2[n-i]) \\
&\quad + \sum_{1\leq i\leq s-1} a^{i-1}u_1[n-i] + \sum_{1\leq i\leq s} a^{i-1}w[n-i] \\
&= (a^s x[n-s] - a^s R_d(y_1[n-s])) + \sum_{1\leq i\leq s} a^{i-1}u_2[n-i] - R_{a^s d}(\sum_{1\leq i\leq s} a^{i-1}u_2[n-i]) \\
&\quad + \sum_{1\leq i\leq s-1} a^{i-1}u_1[n-i] + \sum_{1\leq i\leq s} a^{i-1}w[n-i] \\
&= (a^s(x[n-s] + v_1[n-s] - v_1[n-s]) - a^s R_d(y_1[n-s])) + \sum_{1\leq i\leq s} a^{i-1}u_2[n-i] \\
&\quad - R_{a^s d}(\sum_{1\leq i\leq s} a^{i-1}u_2[n-i]) + \sum_{1\leq i\leq s-1} a^{i-1}u_1[n-i] + \sum_{1\leq i\leq s} a^{i-1}w[n-i] \\
&= (a^s y_1[n-s] - a^s R_d(y_1[n-s])) + \sum_{1\leq i\leq s} a^{i-1}u_2[n-i] - R_{a^s d}(\sum_{1\leq i\leq s} a^{i-1}u_2[n-i]) \\
&\quad + \sum_{1\leq i\leq s-1} a^{i-1}u_1[n-i] + \sum_{1\leq i\leq s} a^{i-1}w[n-i] - a^s v_1[n-s].
\end{aligned}
$$

Here, by Lemmas 4.2 and 4.3 we have

$$a^s y_1[n-s] - a^s R_d(y_1[n-s]) \leq_{df} (a^s d, 0, 0), \tag{4.7}$$

$$\sum_{1\leq i\leq s} a^{i-1}u_2[n-i] - R_{a^s d}(\sum_{1\leq i\leq s} a^{i-1}u_2[n-i]) \leq_{df} (a^s d, 0, 0), \tag{4.8}$$

$$\sum_{1\leq i\leq s} a^{i-1}w[n-i] - a^s v_1[n-s] \sim \mathcal{N}(0, \sum_{1\leq i\leq s} a^{2(i-1)} + a^{2s}\sigma_{v1}^2),$$

$$\sum_{1\leq i\leq s-1} a^{i-1}u_1[n-i] \leq_{df} (\infty, ad + a^2 d + \cdots + a^{s-1}d, 0).$$

The first and second term have a lattice structure. The third and fourth term can be thought as bounded disturbances.

Since

$$\sum_{1\leq i\leq s} a^{2(i-1)} + a^{2s}\sigma_{v1}^2 \leq a^{2(s-1)}\frac{a^2}{a^2-1} + a^{2s}\sigma_{v1}^2$$

$$ad + a^2 d + \cdots + a^{s-1}d = a^{s-1}d(1 + \frac{1}{a} + \cdots + \frac{1}{a^{s-2}}) \leq a^{s-1}d\frac{a}{a-1}$$

by Lemma 4.2, Lemma 4.5 we conclude for all $w_1 \geq 0$

$$\sum_{1\leq i\leq s-1} a^{i-1}u_1[n-i] + \sum_{1\leq i\leq s} a^{i-1}w[n-i] \leq (\infty, a^{s-1}d\frac{a}{a-1} + w_1, 2\cdot Q(\frac{w_1}{2\sqrt{a^{2(s-1)}\frac{a^2}{a^2-1} + a^{2s}\sigma_{v1}^2}})) \tag{4.9}$$

Applying Lemma 4.2 to (4.7), (4.8), (4.9) gives

$$X[n] \leq (a^s d, a^{s-1}d\frac{a}{a-1} + w_1, 2\cdot Q(\frac{w_1}{2\sqrt{a^{2(s-1)}\frac{a^2}{a^2-1} + a^{2s}\sigma_{v1}^2}})).$$

Therefore, we can see that $X[n]$ $(n \geq s)$ has a lattice structure. Then, we will analyze the performance of the estimator of $X[n]$ using Lemma 4.6.

First, for $n \geq s$ we have the following inequality.

$$\begin{aligned}
&\mathbb{E}[(X[n] - Q_{a^s d}(X[n]))^2] \\
&= \mathbb{E}[(x[n] - R_{a^s d}(\sum_{1\leq i\leq s} a^{i-1}u_2[n-i]) - Q_{a^s d}(x[n] - R_{a^s d}(\sum_{1\leq i\leq s} a^{i-1}u_2[n-i])))^2] \\
&= \mathbb{E}[(x[n] - R_{a^s d}(\sum_{1\leq i\leq s} a^{i-1}u_2[n-i]) - Q_{a^s d}(\sum_{1\leq i\leq s} a^{i-1}u_2[n-i]) \\
&\quad - Q_{a^s d}(x[n] - R_{a^s d}(\sum_{1\leq i\leq s} a^{i-1}u_2[n-i]) - Q_{a^s d}(\sum_{1\leq i\leq s} a^{i-1}u_2[n-i])))^2] \\
&= \mathbb{E}[(x[n] - \sum_{1\leq i\leq s} a^{i-1}u_2[n-i] - Q_{a^s d}(x[n] - \sum_{1\leq i\leq s} a^{i-1}u_2[n-i]))^2] \\
&= \mathbb{E}[(\sum_{1\leq i\leq s-1} a^{i-1}u_1[n-i] + \sum_{1\leq i\leq s} a^{i-1}w[n-i] - a^s v_1[n-s] \\
&\quad - Q_{a^s d}(\sum_{1\leq i\leq s-1} a^{i-1}u_1[n-i] + \sum_{1\leq i\leq s} a^{i-1}w[n-i] - a^s v_1[n-s]))^2] \\
&\leq \mathbb{E}[(\sum_{1\leq i\leq s-1} a^{i-1}u_1[n-i] + \sum_{1\leq i\leq s} a^{i-1}w[n-i] - a^s v_1[n-s])^2] \\
&(\because \textit{Lemma } 4.4) \\
&\leq 2\mathbb{E}[(\sum_{1\leq i\leq s-1} a^{i-1}u_1[n-i])^2] + 2\mathbb{E}[(\sum_{1\leq i\leq s} a^{i-1}w[n-i] - a^s v_1[n-s])^2] \\
&(\because \textit{Lemma } 4.1) \\
&\leq 2(\sqrt{\mathbb{E}[u_1^2[n-1]]} + \cdots + \sqrt{a^{2(s-2)}\mathbb{E}[u_1^2[n-s+1]]})^2 + 2\mathbb{E}[(\sum_{1\leq i\leq s} a^{i-1}w[n-1] - a^s v_1[n-s])^2] \\
&(\because \textit{Lemma } 4.1) \\
&\leq 2(\frac{ad}{2})^2 a^{2(s-2)}(\frac{1}{1-\frac{1}{a}})^2 + 2a^{2(s-1)}(\frac{1}{1-\frac{1}{a^2}}) + 2a^{2s}\sigma_{v1}^2 \\
&(\because \text{Definition of } u_1[n]) \\
&= a^{2(s-1)}(2(\frac{d}{2})^2(\frac{1}{1-\frac{1}{a}})^2 + 2(\frac{1}{1-\frac{1}{a^2}}) + 2a^2\sigma_{v1}^2)
\end{aligned}$$

Therefore, by Lemma 4.6 we can bound the estimation error as follows.

$$\begin{aligned}
&\mathbb{E}[(X[n]-Q_{a^sd}(Y_2[n]))^2]\\
&=\mathbb{E}[(X[n]-Q_{a^sd}(X[n]+v_2[n]))^2]\\
&\leq a^{2(s-1)}(2(\frac{d}{2})^2(\frac{1}{1-\frac{1}{a}})^2+2(\frac{1}{1-\frac{1}{a^2}})+2a^2\sigma_{v1}^2)\\
&+(a^sd+\frac{a^{s-1}d\frac{a}{a-1}+w_1}{2})^2 2Q(\frac{a^sd-(a^{s-1}d\frac{a}{a-1}+w_1)}{2\sigma_{v2}})\\
&+(2a^sd+\frac{a^{s-1}d\frac{a}{a-1}+w_1}{2})^2 2Q(\frac{3a^sd-(a^{s-1}d\frac{a}{a-1}+w_1)}{2\sigma_{v2}})+\cdots\\
&+2Q(\frac{w_1}{2\sqrt{a^{2(s-1)}\frac{a^2}{a^2-1}+a^{2s}\sigma_{v1}^2}})((a^sd+\frac{a^sd}{2})^2+(2a^sd+\frac{a^sd}{2})^2 2Q(\frac{a^sd}{\sigma_{v2}})+(3a^sd+\frac{a^sd}{2})^2 2Q(\frac{2a^sd}{\sigma_{v2}})+\cdots)
\end{aligned}$$

Finally, by plugging the above equation into (4.6) we conclude for all $n \geq s$,

$$\begin{aligned}
\mathbb{E}[x^2[n+1]] &= \mathbb{E}[(a(X[n]-Q_{a^sd}(Y_2[n]))+u_1[n]+w[n])^2]\\
&\leq 2\mathbb{E}[(a(X[n]-Q_{a^sd}(Y_2[n])))^2]+2\mathbb{E}[u_1^2[n]]+\mathbb{E}[w^2[n]]\\
&\leq 2a^{2s}(2(\frac{d}{2})^2(\frac{1}{1-\frac{1}{a}})^2+2(\frac{1}{1-\frac{1}{a^2}})+2a^2\sigma_{v1}^2)\\
&+2a^2(a^sd+\frac{a^{s-1}d\frac{a}{a-1}+w_1}{2})^2 2Q(\frac{a^sd-(a^{s-1}d\frac{a}{a-1}+w_1)}{2\sigma_{v2}})\\
&+2a^2(2a^sd+\frac{a^{s-1}d\frac{a}{a-1}+w_1}{2})^2 2Q(\frac{3a^sd-(a^{s-1}d\frac{a}{a-1}+w_1)}{2\sigma_{v2}})+\cdots\\
&+4a^2Q(\frac{w_1}{2\sqrt{a^{2(s-1)}\frac{a^2}{a^2-1}+a^{2s}\sigma_{v1}^2}})((a^sd+\frac{a^sd}{2})^2+(2a^sd+\frac{a^sd}{2})^2 2Q(\frac{a^sd}{\sigma_{v2}})\\
&+(3a^sd+\frac{a^sd}{2})^2 2Q(\frac{2a^sd}{\sigma_{v2}})+\cdots)\\
&+2(a^2(\frac{d}{2})^2)+1 \qquad (4.10)
\end{aligned}$$

Moreover, by (4.6), $\mathbb{E}[x^2[n]]$ is bounded for any $n < s$. Therefore, the L.H.S. of (4.10) is an upper bound on $D(P_1, P_2)$.

For all n, we also have

$$\mathbb{E}[u_1^2[n]] \leq \frac{a^2d^2}{4} \qquad (4.11)$$

which is an upper bound on P_1.

Before we bound $\mathbb{E}[u_2^2[n]]$, we first notice that by Lemma 4.1,

$$\begin{aligned}
&\mathbb{E}[(Q_{a^sd}(y_2[n]-R_{a^sd}(\sum_{1\le i\le s} a^{i-1}u_2[n-1])))^2] \\
&=\mathbb{E}[(y_2[n]-R_{a^sd}(\sum_{1\le i\le s} a^{i-1}u_2[n-1]-R_{a^sd}(y_2[n]-R_{a^sd}(\sum_{1\le i\le s} a^{i-1}u_2[n-1])))^2] \\
&\le 2\mathbb{E}[(y_2[n]-R_{a^sd}(\sum_{1\le i\le s} a^{i-1}u_2[n-1]))^2]+2(\frac{a^sd}{2})^2 \\
&=2\mathbb{E}[(x[n]-R_{a^sd}(\sum_{1\le i\le s} a^{i-1}u_2[n-1]))^2]+2\sigma_{v2}^2+2(\frac{a^sd}{2})^2 \\
&\le 4\mathbb{E}[x^2[n]]+4(\frac{a^sd}{2})^2+2\sigma_{v2}^2+2(\frac{a^sd}{2})^2
\end{aligned}$$

Therefore, for all n,

$$\begin{aligned}
\mathbb{E}[u_2^2[n]] &= a^2\mathbb{E}[(Q_{a^sd}(y_2[n]-R_{a^sd}(\sum_{1\le i\le s} a^{i-1}u_2[n-1]))+R_{a^sd}(\sum_{1\le i\le s} a^{i-1}u_2[n-i]))^2] \\
&\le a^2(2\mathbb{E}[(Q_{a^sd}(y_2[n]-R_{a^sd}(\sum_{1\le i\le s} a^{i-1}u_2[n-1])))^2]+2(\frac{a^sd}{2})^2) \\
&\le a^2(8\mathbb{E}[x^2[n]]+8(\frac{a^sd}{2})^2+4\sigma_{v2}^2+4(\frac{a^sd}{2})^2+2(\frac{a^sd}{2})^2) \\
&\le 8a^2\mathbb{E}[x^2[n]]+\frac{7}{2}a^{2(s+1)}d^2+4a^2\sigma_{v2}^2 \qquad (4.12)
\end{aligned}$$

which gives an upper bound on P_2. Therefore, by (4.10), (4.11), (4.12) the lemma is proved. □

4.6 Proofs and Proof Ideas: Lower bound on the optimal cost

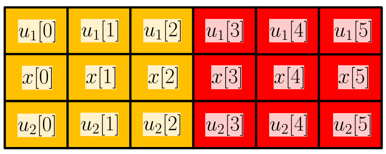


Figure 4.14: Naive truncation idea to divide an infinite-horizon problem to finite-horizon subproblems. This idea fails to give a constant-ratio lower bound.

In this section, we will study the lower bound on the optimal cost and understand why it is impossible to outperform the proposed strategies by an arbitrary factor. In Section 4.5, we discussed the relationship between the infinite-horizon problem of this chapter and finite-horizon problems. The first idea for the lower bound is to make this idea formal, i.e. dividing the infinite-horizon problem into a sequence of finite-horizon problems.

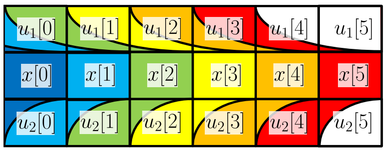


Figure 4.15: Geometric Slicing idea to divide an infinite-horizon problem to finite-horizon sub-problems. This idea successfully gives a lower bound tight to within a constant ratio.

4.6.1 Geometric Slicing of Infinite-Horizon Problems

Let's say we want to divide the infinite-horizon problem into sub-problems with time-horizon 3. A naive way of dividing the problem is truncation, which is pictorially described in Figure 4.14. The total cost $\sum_{i=1}^{N} q\mathbb{E}[x^2[n]] + r_1\mathbb{E}[u_1^2[n]] + r_2\mathbb{E}[u_2^2[n]]$ can be divided into 3-time-horizon problems. The first problem is minimizing $q\mathbb{E}[x^2[0] + x^2[1] + x^2[2]] + r_1\mathbb{E}[u_1^2[0] + u_1^2[1] + u_1^2[2]] + r_2\mathbb{E}[u_2^2[0] + u_2^2[1] + u_2^2[2]]$. The second problem is minimizing $q\mathbb{E}[x^2[3] + x^2[4] + x^2[5]] + r_1\mathbb{E}[u_1^2[3] + u_1^2[4] + u_1^2[5]] + r_2\mathbb{E}[u_2^2[3] + u_2^2[4] + u_2^2[5]]$, and so on. However, in this approach we can find only $\frac{N}{3}$ sub-problems out of N times, which turns out not to be enough to prove constant-ratio optimality.

The main reason why the truncation idea gives too loose a bound is that in order to decouple the sub-problems from each other, we have to start each sub-problem with initial state 0 because that is the best possible initial state. But then, we have to wait long enough until the state disturbance $w[n]$ is amplified enough. So, in each sub-problem, the state cost at the final time step is the only one that is large enough. We end up penalizing the state only for $\frac{N}{3}$-time steps, while the actual cost penalizes the state for N-time steps. In general, if we truncate the problem to s-time-horizon problems, the resulting bound will be loose by a factor of s. In fact, to find a matching lower bound for the s-stage signaling strategy, we have to divide the infinite-horizon problem into s-time-horizon problems. Therefore, the lower bounds based on the truncation idea will be too loose as s goes to infinity.

The idea of 'geometric slicing' solves this by introducing interlocking sub-problems and penalizing the state at every time step. Figure 4.15 shows the idea pictorially. For example, we can slice the problem to 3-time horizon problems as follows. The first problem is minimizing $q\mathbb{E}[x^2[2]] + r_1\mathbb{E}[\frac{1}{2}u_1^2[0] + \frac{1}{4}u_1^2[1] + \frac{1}{8}u_1^2[2]] + r_2\mathbb{E}[\frac{1}{2}u_2^2[1] + \frac{1}{4}u_2^2[2]]$. The second problem is minimizing $q\mathbb{E}[x^2[3]] + r_1\mathbb{E}[\frac{1}{2}u_1^2[1] + \frac{1}{4}u_1^2[2] + \frac{1}{8}u_1^2[3]] + r_2\mathbb{E}[\frac{1}{2}u_2^2[2] + \frac{1}{4}u_2^2[3]]$, and so on. Here, notice that $u_1^2[1]$ shows up in both problems but it does not cause any difficulty since the weights form a geometric sequence and the sum is less than 1. Therefore, we are slicing the problem using geometric sequences, and that is where the name of the idea come from. In this way, we can extract N sub-problems out of an N-time-horizon problem. The sub-problems can be formally written as follows.

Problem G (Geometrically-Sliced Finite-horizon LQG problem for Problem B)**.** *Let the system* *equations, the problem parameters, the underlying random variables, and the restrictions on the controllers be given exactly the same as Problem B. However, now the control objective is for given* $0 < \alpha < 1$, $k, k_1, k_2 \in \mathbb{N}$ $(k_1 \leq k, k_2 \leq k)$*, minimizing the finite-horizon cost*

$$\inf_{u_1,u_2} q\mathbb{E}[x^2[k]] + r_1(1-\alpha)(\sum_{k_1 \leq i \leq k-1} \alpha^{i-k_1}\mathbb{E}[u_1^2[i]]) + r_2(1-\alpha)(\sum_{k_2 \leq i \leq k-1} \alpha^{i-k_2}\mathbb{E}[u_2^2[i]]).$$

Even if the system can run for infinite time, the cost terminates after the time step k. Therefore, this problem is effectively a finite-horizon problem. The next lemma shows the cost of this finite-horizon problem is a lower bound to the original infinite-horizon cost of Problem B.

Lemma 4.8. *Let the system equations, the problem parameters, the underlying random variables, and the restrictions on the controllers be given as in Problem B. When* $\sigma_0^2 = 0$*, for all* $0 < \alpha < 1$, $k, k_1, k_2 \in \mathbb{N}$ $(k_1 \leq k, k_2 \leq k)$*, the infinite-horizon cost of Problem B is lower bounded by the finite-horizon cost of Problem G, i.e.*

$$\begin{aligned}&\inf_{u_1,u_2} \limsup_{N\to\infty} \frac{1}{N} \sum_{0 \leq n \leq N-1} (q\mathbb{E}[x^2[n]] + r_1\mathbb{E}[u_1^2[n]] + r_2\mathbb{E}[u_2^2[n]]) \\ &\geq \inf_{u_1,u_2} q\mathbb{E}[x^2[k]] + r_1(1-\alpha)(\sum_{k_1 \leq i \leq k-1} \alpha^{i-k_1}\mathbb{E}[u_1^2[i]]) + r_2(1-\alpha)(\sum_{k_2 \leq i \leq k-1} \alpha^{i-k_2}\mathbb{E}[u_2^2[i]]).\end{aligned} \tag{4.13}$$

Here, when $k_1 = k$ *or* $k_2 = k$ *the second or third term in the lower bound vanishes.*

Furthermore, both costs are increasing functions of σ_0^2 *and when* $\sigma_0^2 = 0$, $u_1[0] = 0$ *and* $u_2[0] = 0$ *are optimal for both.*

Proof. Let's first prove that for all finite-horizon and infinite horizon problems, the average cost is an increasing function in σ_0^2.

Proposition 4.6. *Let* $x'[0]$ *and* $x''[0]$ *be independent random variables, and* $x'[0]$ *has zero mean. Consider two systems where the system equations are given by Problem B. However, the initial state of the first system is* $x'[0] + x''[0]$ *while the initial state of the second system is* $x'[n]$*. Except for the initial states, both systems have the same underlying random variables* $w[n]$, $v_1[n]$, $v_2[n]$ *as those in Problem B. We denote the variables of the first system as* $x[n]$, $u_i[n]$, $y_i[n]$*, and those of the second system as* $\bar{x}[n]$, $\bar{u}_i[n]$, $\bar{y}_i[n]$*. Then, the following inequality is true.*

$$\begin{aligned}&\inf_{u_1,u_2} \frac{1}{N} \sum_{0 \leq n \leq N-1} (q\mathbb{E}[x^2[n]] + r_1\mathbb{E}[u_1^2[n]] + r_2\mathbb{E}[u_2^2[n]]) \\ &\geq \inf_{\bar{u}_1,\bar{u}_2} \frac{1}{N} \sum_{0 \leq n \leq N-1} (q\mathbb{E}[\bar{x}^2[n]] + r_1\mathbb{E}[\bar{u}_1^2[n]] + r_2\mathbb{E}[\bar{u}_2^2[n]]).\end{aligned}$$

Proof. Since both systems are coupled with each other except the initial state, we will reduce the first system to the second system by giving $x''[0]$ as side-information. Define L_g as the set of strategies for the first system which depend on its own observations and $x''[0]$, i.e. $L_g := \{(u_1[n], u_2[n]) : u_1[n] = f_{1,n}(y_1[0], \cdots, y_1[n], x''[0]), u_2[n] = f_{2,n}(y_2[0], \cdots, y_2[n], x''[0])\}$. Likewise, define L'_g as the set of strategies for the second system which depend on its own observations and $x''[0]$, i.e. $L'_g := \{(\bar{u}_1[n], \bar{u}_2[n]) : \bar{u}_1[n] = f'_{1,n}(\bar{y}_1[0], \cdots, \bar{y}_1[n], x''[0]), f'_{2,n}(\bar{y}_1[0], \cdots, \bar{y}_1[n], x''[0])\}$.

Further, define $u'_i[n] := u_i[n] - \mathbb{E}[u_i[n]|x''[0]]$ and $u''_i[n] := \mathbb{E}[u_i[n]|x''[0]]$. Then, we can lower bound the average cost as follows.

$$\inf_{u_1,u_2} \frac{1}{N} \sum_{0\leq n\leq N-1} (q\mathbb{E}[x^2[n]] + r_1\mathbb{E}[u_1^2[n]] + r_2\mathbb{E}[u_2^2[n]])$$

$$\overset{(A)}{\geq} \inf_{u_1,u_2\in L_g} \frac{1}{N} \sum_{0\leq n\leq N-1} (q\mathbb{E}[x^2[n]] + r_1\mathbb{E}[u_1^2[n]] + r_2\mathbb{E}[u_2^2[n]])$$

$$\overset{(B)}{=} \inf_{u_1,u_2\in L_g} \frac{1}{N} \sum_{0\leq n\leq N-1} (q\mathbb{E}[(a^n x[0] + a^{n-1}w[0] + \cdots + w[n-1]$$

$$+ a^{n-1}u_1[0] + \cdots + u_1[n-1] + a^{n-1}u_2[0] + \cdots + u_2[n-1])^2]$$

$$+ r_1\mathbb{E}[u_1^2[n]] + r_2\mathbb{E}[u_2^2[n]])$$

$$\overset{(C)}{=} \inf_{u_1,u_2\in L_g} \frac{1}{N} \sum_{0\leq n\leq N-1} q\mathbb{E}[(a^n x'[0] + a^{n-1}w[0] + \cdots + w[n-1]$$

$$+ a^{n-1}u_1'[0] + \cdots + u_1'[n-1] + a^{n-1}u_2'[0] + \cdots + u_2'[n-1]$$

$$+ a^n x''[0] + a^{n-1}u_1''[0] + \cdots + u_1''[n-1] + a^{n-1}u_2''[0] + \cdots + u_2''[n-1])^2]$$

$$+ r_1\mathbb{E}[(u_1'[n] + u_1''[n])^2] + r_2\mathbb{E}[(u_2'[n] + u_2''[n])^2]$$

$$\overset{(D)}{=} \inf_{u_1,u_2\in L_g} \frac{1}{N} \sum_{0\leq n\leq N-1} q\mathbb{E}[(a^n x'[0] + a^{n-1}w[0] + \cdots + w[n-1]$$

$$+ a^{n-1}u_1'[0] + \cdots + u_1'[n-1] + a^{n-1}u_2'[0] + \cdots + u_2'[n-1])^2]$$

$$+ q\mathbb{E}[(a^n x''[0] + a^{n-1}u_1''[0] + \cdots + u_1''[n-1] + a^{n-1}u_2''[0] + \cdots + u_2''[n-1])^2]$$

$$+ r_1\mathbb{E}[u_1'[n]^2] + r_1\mathbb{E}[u_1''[n]^2] + r_2\mathbb{E}[u_2'[n]^2] + r_2\mathbb{E}[u_2''[n]^2]$$

$$\geq \inf_{u_1,u_2\in L_g} \frac{1}{N} \sum_{0\leq n\leq N-1} q\mathbb{E}[(a^n x'[0] + a^{n-1}w[0] + \cdots + w[n-1]$$

$$+ a^{n-1}u_1'[0] + \cdots + u_1'[n-1] + a^{n-1}u_2'[0] + \cdots + u_2'[n-1])^2]$$

$$+ r_1\mathbb{E}[u_1'[n]^2] + r_2\mathbb{E}[u_2'[n]^2]$$

$$\overset{(E)}{\geq} \inf_{\bar{u}_1,\bar{u}_2\in L_g'} \frac{1}{N} \sum_{0\leq n\leq N-1} (q\mathbb{E}[\bar{x}^2[n]] + r_1\mathbb{E}[\bar{u}_1^2[n]] + r_2\mathbb{E}[\bar{u}_2^2[n]])$$

$$= \inf_{\bar{u}_1,\bar{u}_2\in L_g'} \frac{1}{N}\mathbb{E}[\sum_{0\leq n\leq N-1} (q\mathbb{E}[\bar{x}^2[n]] + r_1\mathbb{E}[\bar{u}_1^2[n]] + r_2\mathbb{E}[\bar{u}_2^2[n]])|x''[0]]$$

$$\geq \inf_{x'} \inf_{\bar{u}_1,\bar{u}_2\in L_g'} \frac{1}{N}\mathbb{E}[\sum_{0\leq n\leq N-1} (q\mathbb{E}[\bar{x}^2[n]] + r_1\mathbb{E}[\bar{u}_1^2[n]] + r_2\mathbb{E}[\bar{u}_2^2[n]])|x''[0] = x']$$

$$\overset{(F)}{=} \inf_{\bar{u}_1,\bar{u}_2\in L} \frac{1}{N} \sum_{0\leq n\leq N-1} (q\mathbb{E}[\bar{x}^2[n]] + r_1\mathbb{E}[\bar{u}_1^2[n]] + r_2\mathbb{E}[\bar{u}_2^2[n]])$$

(A): $L \subseteq L_g$.

(B): By the system dynamics of (4.2).

(C): Definitions of $x'[0]$, $x''[0]$, $u_i'[n]$, $u_i''[n]$.

(D): Since $x'[0], w[n]$ are zero mean and independent from $x''[0]$, they are orthogonal to $x''[0]$, $u_1''[n]$, $u_2''[n]$. Moreover, by definition, $u_1'[n]$, $u_2'[n]$ are orthogonal to $x''[0]$, $u_1''[n]$, $u_2''[n]$.

(E): To justify this, we will show (by induction) that for all n, $y_i[0], \cdots y_i[n], x''[0]$ are functions of $\bar{y}_i[0], \cdots \bar{y}_i[n], x''[0]$. Therefore, there exists $\bar{u}_1[n], \bar{u}_2[n]$ such that $\bar{u}_1[n] = u'_1[n]$, $\bar{u}_2[n] = u'_2[n]$.

First, when $n = 1$, the claim is obvious since $y_i[0] = \bar{y}_i[0] + x''[0]$. Thus, $y_i[0], x''[0]$ are functions of $\bar{y}_i[0], x''[0]$. Moreover, since $u'_i[n]$ are functions of $y_i[0]$ and $x''[0]$, we can find $\bar{u}_i[n]$ such that $\bar{u}_i[n] = u'_i[n]$.

Let's say the claim holds until $n-1$. Then, we have

$$\begin{aligned} y_i[n] =& a^n x'[0] + a^{n-1}w[0] + \cdots + w[n-1] \\ &+ a^{n-1}u'_1[0] + \cdots + u'_1[n-1] \\ &+ a^{n-1}u'_2[0] + \cdots + u'_2[n-1] + v_i[n] + g(x''[0]) \\ =& a^n x'[0] + a^{n-1}w[0] + \cdots + w[n-1] \\ &+ a^{n-1}\bar{u}_1[0] + \cdots + \bar{u}_1[n-1] \\ &+ a^{n-1}\bar{u}_2[0] + \cdots + \bar{u}_2[n-1] + v_i[n] + g(x''[0]) \\ =& \bar{y}_i[n] + g(x''[0]) \end{aligned}$$

where $g(x''[0]) := a^n x''[0] + a^{n-1}\mathbb{E}[u_1[0]|x''[0]] + \cdots + \mathbb{E}[u_1[n-1]|x''[0]] + a^{n-1}\mathbb{E}[u_2[0]|x''[0]] + \cdots + \mathbb{E}[u_2[n-1]|x''[0]]$, and the send equality comes from the induction hypothesis. Therefore, $y_i[n]$ is a function of $\bar{y}_i[n], x''[0]$, and we can find $\bar{u}_i[n]$ such that $\bar{u}_i[n] = u'_i[n]$. This proves the claim by induction.

(F): Since in L'_g the strategies can depend on $x''[0]$.

Therefore, the proposition is true. □

Here, we can notice that the proof holds for all quadratic costs. Therefore, by setting $x'[0] \sim \mathcal{N}(0, \sigma_0'^2)$ and $x''[0] \sim \mathcal{N}(0, \sigma_0''^2)$, we can prove the costs in (4.13) are increasing functions on σ_0^2. We can also easily justify that when $x_0[0] = 0$, $u_1[0] = u_2[0] = 0$ is the optimal input by symmetry.

Then, let's prove the inequality of (4.13).

$$
\begin{aligned}
&\inf_{u_1,u_2} \limsup_{N\to\infty} \frac{1}{N} \sum_{0\le n\le N-1} (q\mathbb{E}[x^2[n]] + r_1\mathbb{E}[u_1^2[n]] + r_2\mathbb{E}[u_2^2[n]]) \\
&\overset{(a)}{\ge} \limsup_{N\to\infty} \inf_{u_1,u_2} \frac{1}{N} \sum_{0\le n\le N-1} (q\mathbb{E}[x^2[n]] + r_1\mathbb{E}[u_1^2[n]] + r_2\mathbb{E}[u_2^2[n]]) \\
&\overset{(b)}{\ge} \limsup_{N\to\infty} \inf_{u_1,u_2} \frac{1}{N} (q\mathbb{E}[x^2[k]] + r_1(1-\alpha)(\sum_{k_1\le i\le k-1} \alpha^{i-k_1}\mathbb{E}[u_1^2[i]]) + r_2(1-\alpha)(\sum_{k_2\le i\le k-1} \alpha^{i-k_2}\mathbb{E}[u_2^2[i]]) \\
&\quad + q\mathbb{E}[x^2[k+1]] \\
&\quad + r_1(1-\alpha)(\sum_{k_1\le i\le k-1} \alpha^{i-k_1}\mathbb{E}[u_1^2[i+1]]) + r_2(1-\alpha)(\sum_{k_2\le i\le k-1} \alpha^{i-k_2}\mathbb{E}[u_2^2[i+1]]) + \cdots \\
&\quad + q\mathbb{E}[x^2[N-1]] \\
&\quad + r_1(1-\alpha)(\sum_{k_1\le i\le k-1} \alpha^{i-k_1}\mathbb{E}[u_1^2[i+N-k-1]]) + r_2(1-\alpha)(\sum_{k_2\le i\le k-1} \alpha^{i-k_2}\mathbb{E}[u_2^2[i+N-k-1]]) \\
&\overset{(c)}{\ge} \limsup_{N\to\infty} \frac{1}{N} (\inf_{u_1,u_2} q\mathbb{E}[x^2[k]] + r_1(1-\alpha)(\sum_{k_1\le i\le k-1} \alpha^{i-k_1}\mathbb{E}[u_1^2[i]]) + r_2(1-\alpha)(\sum_{k_2\le i\le k-1} \alpha^{i-k_2}\mathbb{E}[u_2^2[i]]) \\
&\quad + \inf_{u_1,u_2} q\mathbb{E}[x^2[k+1]] \\
&\quad + r_1(1-\alpha)(\sum_{k_1\le i\le k-1} \alpha^{i-k_1}\mathbb{E}[u_1^2[i+1]]) + r_2(1-\alpha)(\sum_{k_2\le i\le k-1} \alpha^{i-k_2}\mathbb{E}[u_2^2[i+1]]) + \cdots \\
&\quad + \inf_{u_1,u_2} q\mathbb{E}[x^2[N-1]] \\
&\quad + r_1(1-\alpha)(\sum_{k_1\le i\le k-1} \alpha^{i-k_1}\mathbb{E}[u_1^2[i+N-k-1]]) + r_2(1-\alpha)(\sum_{k_2\le i\le k-1} \alpha^{i-k_2}\mathbb{E}[u_2^2[i+N-k-1]]) \qquad (4.14) \\
&\overset{(d)}{\ge} \limsup_{N\to\infty} \frac{N-k}{N} (\inf_{u_1,u_2} q\mathbb{E}[x^2[k]] + r_1(1-\alpha)(\sum_{k_1\le i\le k-1} \alpha^{i-k_1}\mathbb{E}[u_1^2[i]]) + r_2(1-\alpha)(\sum_{k_2\le i\le k-1} \alpha^{i-k_2}\mathbb{E}[u_2^2[i]]) \\
&\overset{(e)}{=} \inf_{u_1,u_2} q\mathbb{E}[x^2[k]] + r_1(1-\alpha)(\sum_{k_1\le i\le k-1} \alpha^{i-k_1}\mathbb{E}[u_1^2[i]]) + r_2(1-\alpha)(\sum_{k_2\le i\le k-1} \alpha^{i-k_2}\mathbb{E}[u_2^2[i]]).
\end{aligned}
$$

(a): $\inf\sup \ge \sup\inf$.

(b): We can easily check that the sum of the weight for each input cost, $\mathbb{E}[u_1^2[n]]$ or $\mathbb{E}[u_2^2[n]]$ is less than $(1-\alpha)(1+\alpha+\alpha^2+\cdots)$ which is 1.

(c): $\inf_x f(x) + g(x) \ge \inf_x f(x) + \inf_{x'} g(x')$.

(d): The second minimization problem in (4.14) can be thought as a one-time-step shift of the first minimization problem, i.e. $x[1]$ of the second problem corresponds to the initial state $x[0]=0$ of the first problem. Therefore, by putting $x'[0]=0$ and $x''[0]=x[1]$ in Proposition 4.6, the first problem's cost is smaller than the second problem's cost. Likewise, we can prove that the first problem's cost is a lower bound for all other problems' cost.

(e): $\limsup_{N\to\infty} \frac{N-k}{N} = 1$. □

Conceptually, this idea of geometric slicing can be thought of as an interesting variant on how discounted dynamic programming [11] is used to study average-cost dynamic programming.

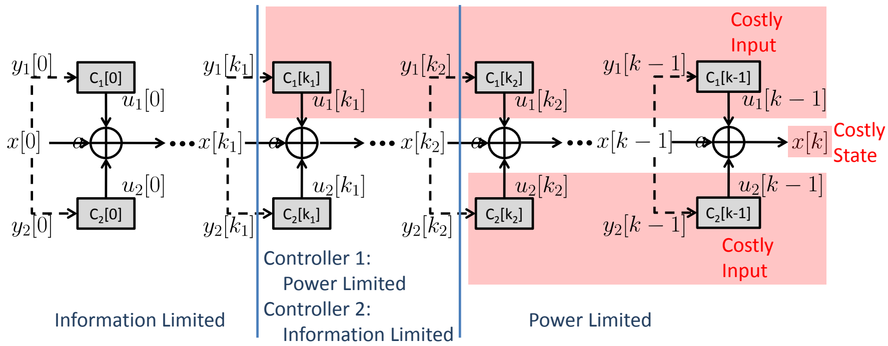


Figure 4.16: The general finite-horizon problem structure which can give a lower bound for s-stage signaling strategies. The problem consists of three time intervals. In the first time interval, both controllers are information-limited. In the second time interval, the first contoller is power-limited and the second controller is information-limited. In the third time interval, both controllers are power-limited.

4.6.2 Finite-Horizon LQG Problems: Three-Stage Division

Now, we can divide the infinite-horizon problem to finite-horizon problems. Figure 4.16 shows the finite-horizon problem that gives a lower bound approximately matching with s-stage signaling strategies. As we discussed in Figure 4.8, the resulting problem is not stationary and to tackle this issue we will divide the time-horizon into three intervals: (1) information-limited interval, (2) MIMO Witsenhausen's interval, (3) power-limited interval.

Let's first state the power-distortion tradeoff version of the finite-horizon problem of Problem G.[13]

Problem H (Finite-Horizon LQG problem with discounted power constraints)**.** *Let's consider the same system and parameters as Problem G. But, now the control objective is minimizing the final state disturbance $D_F(P_1, P_2)$ for given input power constraints $P_1, P_2 \in \mathbb{R}^+$. In other words, we*

[13]This is not a finite-hirizon version of Problem F.

solve

$$\begin{aligned} D_F(P_1,P_2) = &\inf_{u_1,u_2} \mathbb{E}[x^2[k]] \\ s.t. &\sum_{k_1 \le i \le k-1} \alpha^{i-k_1}\mathbb{E}[u_1^2[i]] \le P_1 \\ &\sum_{k_2 \le i \le k-1} \alpha^{i-k_2}\mathbb{E}[u_2^2[i]] \le P_2. \end{aligned}$$

Here we can see four parameters that characterize the problem: σ_{v1}^2, P_1, σ_{v2}^2, P_2. The importance of these parameters becomes different depending on which interval they lie in.

The information-limited interval —which corresponds to the time steps between 0 and k_1 in Problem H and Figure 4.16— is introduced to handle the case when σ_{v1}^2 is large. Since $\sigma_{v2}^2 \ge \sigma_{v1}^2$, in this interval both controllers have very noisy observations and we can allow arbitrarily large power to both controllers. In fact, in Figure 4.16 we can see in this interval both controllers do not have any input costs. Therefore, the important parameters are σ_{v1}^2 and σ_{v2}^2. It turns out that the cost of the centralized controller (with access to both noisy observations $y_1[n]$ and $y_2[n]$) gives a reasonable bound. Essentially, what this interval is doing is waiting until the variance of the state disturbances grows enough — to be around σ_{v1}^2 up to scaling.

On the other hand, the power-limited interval — which corresponds to the time steps between k_2 and k in Problem H and Figure 4.16 — is introduced to handle cases when both controllers do not have enough power to stabilize the system. Therefore, in this interval the important parameters are P_1 and P_2. We will even give a perfect observation of $x[n]$ to both controllers by setting $\sigma_{v1}^2 = 0$ and $\sigma_{v2}^2 = 0$. In this interval, we will keep running the system by making k arbitrarily large, and prove that $\mathbb{E}[x^2[k]]$ must diverge to infinity given that the previous interval ended up with a too large $x[k_2]$.

Between these two intervals — the time steps between k_1 and k_2 in Problem H and Figure 4.16 — each controller faces a different situation. The first controller has enough information about the state but it does not have enough power. The second controller has enough power but it does not have enough information. Therefore, the important parameters of this interval are P_1 and σ_{v2}^2. So, we will allow a perfect observation to the first controller by setting $\sigma_{v1}^2 = 0$ and infinite power to the second controller by setting $P_2 = \infty$. In other words, the first controller is power limited and the second controller is information limited. This situation is exactly the same as that of Witsenhausen's counterexample which we discussed in Section 4.3.1. Therefore, we will call this interval an s-stage MIMO Witsenhausen's interval and discuss it in Section 4.6.3 in more detail.

Let's convert these ideas into formal proofs. As we mentioned, we will bound the cost in the information-limited interval by analyzing a centralized controller with both observations $y_1[n]$ and $y_2[n]$ when there is only initial disturbance $w[0]$.

Lemma 4.9. *Let* $w[0] \sim \mathcal{N}(0,1)$, $v_1[n] \sim \mathcal{N}(0,\sigma_{v1}^2)$, $v_2[n] \sim \mathcal{N}(0,\sigma_{v2}^2)$ *be independent Gaussian* *random variables. Let*

$$y_1[n] = a^{n-1}w[0] + v_1[n],$$
$$y_2[n] = a^{n-1}w[0] + v_2[n].$$

Then,

$$\mathbb{E}[(a^{k-1}w[0] - \mathbb{E}[a^{k-1}w[0]|y_1[1:k_1], y_2[1:k_1]])^2] = \frac{a^{2(k-1)}\sigma_{v1}^2}{(1+\frac{\sigma_{v1}^2}{\sigma_{v2}^2})(\frac{a^{2(k_1-1)}(1-a^{-2k_1})}{1-a^{-2}}) + \sigma_{v1}^2}.$$

Proof. Notice that

$$y_1[n] = a^{n-1}w[0] + v_1[n]$$
$$\frac{\sigma_{v1}}{\sigma_{v2}}y_2[n] = \frac{\sigma_{v1}}{\sigma_{v2}}a^{n-1}w[0] + \frac{\sigma_{v1}}{\sigma_{v2}}v_2[n]$$

Since maximum-ratio combining is a sufficient statistic (See [99] for instance), the sufficient statistic y_s of $y_1[1:k_1-1], y_2$ for estimating $w[0]$ is given as:

$$\begin{aligned} y_s &= \sum_{1\leq n\leq k_1} a^{n-1}y_1[n] + \sum_{1\leq n\leq k_1} \frac{\sigma_{v1}}{\sigma_{v2}}a^{n-1}(\frac{\sigma_{v1}}{\sigma_{v2}}y_2[n]) \\ &= \sum_{1\leq n\leq k_1} a^{n-1}(a^{n-1}w[0] + v_1[n]) + \sum_{1\leq n\leq k_1} \frac{\sigma_{v1}}{\sigma_{v2}}a^{n-1}(\frac{\sigma_{v1}}{\sigma_{v2}}a^{n-1}w[0] + \frac{\sigma_{v1}}{\sigma_{v2}}v_2[n]) \\ &= (\sum_{1\leq n\leq k_1}(a^{2(n-1)} + \frac{\sigma_{v1}^2}{\sigma_{v2}^2}a^{2(n-1)}))w[0] + (\sum_{1\leq n\leq k_1} a^{n-1}v_1[n] + \sum_{1\leq n\leq k_1}\frac{\sigma_{v1}^2}{\sigma_{v2}^2}a^{n-1}v_2[n]) \end{aligned}$$

The estimation error for $a^{k-1}w[0]$ is

$$\mathbb{E}[(a^{k-1}w[0]-\mathbb{E}[a^{k-1}w[0]|y_1[1:k_1],y_2[1:k_1]])^2]$$
$$=\mathbb{E}[(a^{k-1}w[0]-\mathbb{E}[a^{k-1}w[0]|y_s])^2]$$
$$=\mathbb{E}[(a^{k-1}w[0])^2]-\mathbb{E}[a^{k-1}w[0]y_s](\mathbb{E}[y_s^2])^{-1}\mathbb{E}[a^{k-1}w[0]y_s]$$
$$=\frac{\mathbb{E}[(a^{k-1}w[0])^2]\mathbb{E}[y_s^2]-\mathbb{E}[a^{k-1}w[0]y_s]^2}{\mathbb{E}[y_s^2]}$$
$$=\frac{a^{2(k-1)}((\sum_{1\le n\le k_1}(a^{2(n-1)}+\frac{\sigma_{v1}^2}{\sigma_{v2}^2}a^{2(n-1)}))^2+\sum_{1\le n\le k_1}a^{2(n-1)}\sigma_{v1}^2+\sum_{1\le n\le k_1}(\frac{\sigma_{v1}^2}{\sigma_{v2}^2})^2a^{2(n-1)}\sigma_{v2}^2)}{(\sum_{1\le n\le k_1}(a^{2(n-1)}+\frac{\sigma_{v1}^2}{\sigma_{v2}^2}a^{2(n-1)}))^2+\sum_{1\le n\le k_1}a^{2(n-1)}\sigma_{v1}^2+\sum_{1\le n\le k_1}(\frac{\sigma_{v1}^2}{\sigma_{v2}^2})^2a^{2(n-1)}\sigma_{v2}^2}$$

$$-\frac{a^{2(k-1)}(\sum_{1\le n\le k_1}(a^{2(n-1)}+\frac{\sigma_{v1}^2}{\sigma_{v2}^2}a^{2(n-1)}))^2}{(\sum_{1\le n\le k_1}(a^{2(n-1)}+\frac{\sigma_{v1}^2}{\sigma_{v2}^2}a^{2(n-1)}))^2+\sum_{1\le n\le k_1}a^{2(n-1)}\sigma_{v1}^2+\sum_{1\le n\le k_1}(\frac{\sigma_{v1}^2}{\sigma_{v2}^2})^2a^{2(n-1)}\sigma_{v2}^2}$$
$$=\frac{a^{2(k-1)}(\sum_{1\le n\le k_1}a^{2(n-1)}\sigma_{v1}^2+\sum_{1\le n\le k_1}(\frac{\sigma_{v1}^2}{\sigma_{v2}^2})^2a^{2(n-1)}\sigma_{v2}^2)}{(\sum_{1\le n\le k_1}(a^{2(n-1)}+\frac{\sigma_{v1}^2}{\sigma_{v2}^2}a^{2(n-1)}))^2+\sum_{1\le n\le k_1}a^{2(n-1)}\sigma_{v1}^2+\sum_{1\le n\le k_1}(\frac{\sigma_{v1}^2}{\sigma_{v2}^2})^2a^{2(n-1)}\sigma_{v2}^2}$$
$$=\frac{a^{2(k-1)}(\sigma_{v1}^2+\frac{\sigma_{v1}^4}{\sigma_{v2}^2})(\frac{a^{2(k_1-1)}(1-a^{-2k_1})}{1-a^{-2}})}{(1+\frac{\sigma_{v1}^2}{\sigma_{v2}^2})^2(\frac{a^{2(k_1-1)}(1-a^{-2k_1})}{1-a^{-2}})^2+(\sigma_{v1}^2+\frac{\sigma_{v1}^4}{\sigma_{v2}^2})\frac{a^{2(k_1-1)}(1-a^{-2k_1})}{1-a^{-2}}}$$
$$=\frac{a^{2(k-1)}\sigma_{v1}^2}{(1+\frac{\sigma_{v1}^2}{\sigma_{v2}^2})(\frac{a^{2(k_1-1)}(1-a^{-2k_1})}{1-a^{-2}})+\sigma_{v1}^2}.$$

□

To bound the performance in the power-limited interval, we have to bound the influence of control inputs on the state with respect to their power constraints. By expanding $x[n]$ using the system equation (4.2), we can see $x[n]=\sum_{0\le i\le n-1}a^{n-1-i}w[i]+a^{n-1-i}u_1[i]+a^{n-1-i}u_2[i]$. Thus, the terms $\sum_{0\le i\le n-1}a^{n-1-i}u_j[i]$ can be considered as the influence of control inputs on the state. The following Cauchy-Schwarz style inequality bounds the variance of $\sum_{0\le i\le n-1}a^{n-1-i}u_j[i]$ by the power constraint $\sum_{0\le i\le n-1}\alpha^i\mathbb{E}[u_j^2[i]]$ imposed in Problem H.

Lemma 4.10. *For arbitrary random variables X_i, $n\in\mathbb{N}$, $a,b\in\mathbb{R}$ ($|\frac{1}{a^2b}|<1$), we have*

$$\mathbb{E}[(a^{n-1}X_0+a^{n-2}X_1+\cdots+X_{n-1})^2]\le\frac{a^{2(n-1)}(1-(\frac{1}{a^2b})^n)}{1-\frac{1}{a^2b}}(\mathbb{E}[X_0^2]+b\mathbb{E}[X_1^2]+\cdots+b^{n-1}\mathbb{E}[X_{n-1}^2])$$

Proof.

$$
\begin{aligned}
&\mathbb{E}[(a^{n-1}X_0 + a^{n-2}X_1 + \cdots + X_{n-1})^2] \\
&\leq (\sqrt{a^{2(n-1)}\mathbb{E}[X_0^2]} + \sqrt{a^{2(n-2)}\mathbb{E}[X_1^2]} + \cdots + \sqrt{\mathbb{E}[X_{n-1}^2]})^2 \\
&\leq (a^{2(n-1)} + \frac{a^{2(n-2)}}{b} + \cdots + (\frac{1}{b})^{n-1})(\mathbb{E}[X_0^2] + b\mathbb{E}[X_1^2] + \cdots + b^{n-1}\mathbb{E}[X_{n-1}^2]) \\
&= \frac{a^{2(n-1)}(1-(\frac{1}{a^2 b})^n)}{1-\frac{1}{a^2 b}}(\mathbb{E}[X_0^2] + b\mathbb{E}[X_1^2] + \cdots + b^{n-1}\mathbb{E}[X_{n-1}^2])
\end{aligned}
$$

where all inequalities follow from Cauchy-Schwarz. □

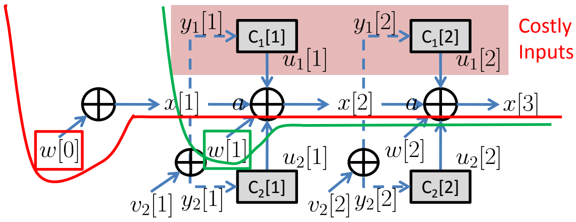


Figure 4.17: Finite-horizon generalized MIMO Witsenhausen's counterexample. This problem gives the matching lower bound to 1-stage signaling.

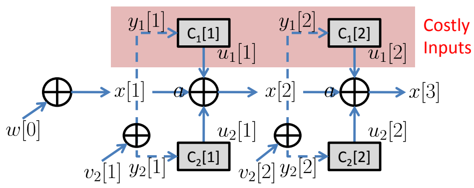


Figure 4.18: The simplified problem that results from Figure 4.17 by cutting the problem across the red line. Unlike the original problem, $w[0]$ is the only disturbance.

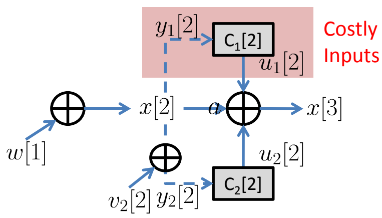
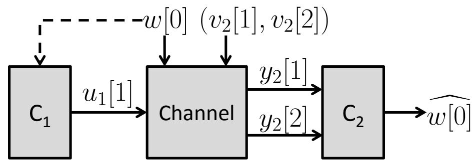


Figure 4.19: The simplified problem that results from Figure 4.17 by cutting the problem across the blue line. Unlike the original problem, $w[1]$ is the only disturbance.



$$\begin{bmatrix} y_2[1] \\ y_2[2] \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ a & 0 & 1 \end{bmatrix} \begin{bmatrix} w[0] \\ v_2[1] \\ v_2[2] \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_1[1]$$

Figure 4.20: A further simplified MIMO communication problem of Figure 4.18. This problem reflects the communication aspect of Figure 4.18.

4.6.3 Tackling the s-stage MIMO Witsenhausen's interval: From multiple disturbances to a single disturbance

Understanding the MIMO Witsenhausen's interval is necessary to find a matching lower bound to s-stage signaling strategies. Let's explicitly consider Problem F with parameters $\sigma_{v1}^2 = 0$ and $P_2 = \infty$ and find the lower bound on $D(P_1, P_2)$ that approximately matches to the 1-stage signaling strategy.

By selecting the parameters $k = 3, k_1 = 1, k_2 = 3, \alpha = \frac{1}{2}$ in Problem H, we have the problem of minimizing $D_F(P_1, P_2) = \mathbb{E}[x^2[3]]$ with the power constraint $(\frac{1}{2}\mathbb{E}[u_1^2[1]] + \frac{1}{4}\mathbb{E}[u_1^2[2]]) \leq P_1$. In the same way as the proof of Lemma 4.8, we can prove that this is a lower bound on $D(P_1, P_2)$.

Figure 4.17 shows the resulting 2-stage finite-horizon problem. As we can see the problem looks similar to Witsenhausen's one in Figure 4.5a. In fact, it can be thought as a multi-stage MIMO (multiple-input multiple-output) Witsenhausen's counterexample. Compared to the original Witsenhausen's counterexample, both controllers have observations and control inputs at every time step, and a new state disturbance $w[n]$ is added at every time step. Since the second controller's input is free, it can be considered as the receiver in a communication problem. From this perspective, the observation $y_2[1]$ can be considered as side-information at the receiver, and the input $u_2[1]$ can be imagined to be feedback from receiver to transmitter.

The first question that we have to answer to take this communication perspective is "What is the relevant message in this communication problem?" Since the only uncertainty of the system is the state disturbance $w[n]$, the answer has to be the disturbance. However, since a new $w[n]$ is added at every time step, we have to find the critically relevant disturbance among them.

To understanding this issue, let's revisit the binary deterministic model of Section 4.3. In Figure 4.2, we can see $x[3]$ corresponds to $00x_{-1}^3 0.x_{-3}^3 x_{-4}^3 \cdots$ in the binary deterministic model. We will divide this binary number into three parts. The first part is the first two bits 00, the second part is the next two bits $x_{-1}^3 0$, and the third part is the remaining bits $x_{-3}^3 x_{-4}^3 \cdots$. If we track back the arrows of Figure 4.2, we can see that these three parts originated from the different disturbances

$w[0]$, $w[1]$, $w[2]$ respectively. Therefore, we can see that $w[2]$ is not a dominating disturbance since its bit level is much smaller than the other parts, and the dominant disturbances for $x[3]$ are $w[0]$ and $w[1]$. We will separate these two disturbances using the cutset idea in information theory.

The first cut gives every disturbance except $w[0]$ as side information to the second controller, i.e. we give $w[1], w[2]$ as side information. Figure 4.18 shows the resulting problem, which is a 2-stage MIMO Witsenhausen's counterexample with only one disturbance at the beginning. Likewise, the second cut gives $w[0], w[2]$ and reserves $w[1]$ inside the cut. Figure 4.19 shows the resulting problem, which is a 1-stage Radner's problem. Both problems are relaxations of the original problem, and any convex sum of their cost is also a lower bound to the cost of the original problem.

We already know how to solve Radner's problem in Figure 4.19. However, the problem in Figure 4.18 is a generalized MIMO Witsenhausen's problem, which is even harder than the original one. The crux of the problem is the dual role of controllers' inputs. The input signals $u_1[n]$ and $u_2[n]$ can be used to cancel the state (control role) and at the same time to send information about their observations (communication role). Therefore, we will simplify the problem by removing the less important role.

The first controller has a perfect observation while its input cost is expensive. Therefore, it is better to use the control inputs to send information about the state. We will essentially remove the control role of the first controller input by using the Cauchy-Schwarz inequality. Meanwhile, the second controller has free input cost but blurry observations. Therefore, it is better to focus on the control role. We will remove the communication role of the second controller input by allowing free freeback from the second controller to the first. Therefore, the first controller reduces to a transmitter and the second controller reduces to a receiver.

Figure 4.20 shows the pure MIMO communication problem we will get after removing the dual roles of the controllers from the problem of Figure 4.19. The first controller knows the exact state $w[0]$ and sends information through the input $u_1[1]$. Thus, the first controller is the transmitter and $u_1[1]$ is the transmitted signal.[14] The second controller estimates the state $w[0]$ based on its observation $y_2[1], y_2[2]$. Therefore, the second controller is the receiver and $y_2[1], y_2[2]$ are the received signals.[15] We will use a simple information-theoretic cutset bound to bound the performance of this communication system, and eventually derive a lower bound approximately matching to the 1-stage signaling strategy.

At this point, one may wonder why we need the lower bound of Figure 4.18 and Figure 4.20 which correspond to zeros in the binary deterministic model. It is because it is not zero in Gaussian real models. Binary deterministic models simplify Gaussian random variables as bounded uniform distributions. This simplification can be justified in an infinite-dimensional relaxation. However, in

[14]Here, $u_1[2]$ cannot send any information to the second controller since communication requires at least one step delay from the transmitter to the receiver.

[15]The second controller can also feedback its observation through $u_2[1]$. However, this effect of feedback is negligible in this case, since the causal feedback information can only affect $u_2[2]$ at the transmitter. However, we will see the effect of feedback later in the more generalized problem of Figure 4.21.

finite dimensions the simplification only approximately holds and the zeros in the binary deterministic model are actually exponentially decreasing small quantities in a Gaussian model. As shown in [37], we will replace $v_2[n]$ of Figure 4.20 by a test channel, adapting ideas of large-deviation theory. The problem of Figure 4.20 gives a non-trivial lower bound that captures the exponentially decreasing small quantities that must occur because of the finite-dimensionality.

In general, we will see an s-stage MIMO Witsenhausen's counterexample in the second time interval of Figure 4.16. Following the same steps as above, we will reduce the problem to a pair of pure communication problems, s-stage and $(s-1)$-stage MIMO state-amplification with feedback.

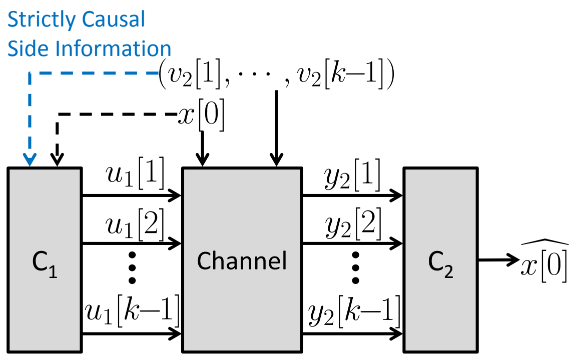


Figure 4.21: s-stage MIMO state-amplification with feedback. This problem reflects the implicit communication aspect in the MIMO Witsenhausen's interval of Figure 4.16. This figure also represents a system diagram of Problem I.

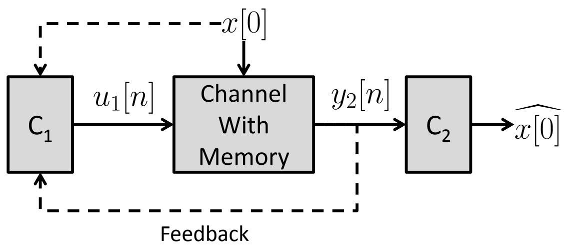


Figure 4.22: An equivalent representation of s-stage MIMO state-amplification with feedback in Figure 4.21. The MIMO channel of Figure 4.21 can be thought as a channel with memory. This figure also represents a system diagram of Problem I.

4.6.4 s-stage MIMO state-amplification with feedback

Figure 4.21 shows the s-stage MIMO state-amplification problem. As we discussed before, the first controller C_1 is the transmitter, and the second controller C_2 is the receiver. The transmitter knows the state $x[0]$ at the initial time and learns the channel noise $v_2[n]$ by causal feedback. The goal of communication is minimizing the estimation error of the state $x[0]$ at the receiver.

Let's formally state the s-stage MIMO state-amplification with feedback problem.

Problem I (s-stage MIMO state-amplification with feedback). *Let the underlying random variables $x[0] \sim \mathcal{N}(0, \sigma_0^2)$ and $v_2[n] \sim \mathcal{N}(0, \sigma_v^2)$ be all independent. These are the source and observation noise respectively. The transmitter's input $u_1[n]$ is a function of $x[0]$ and $v_2[1], \cdots, v_2[n-1]$, i.e.*

$$
\begin{aligned}
&u_1[1] = f_1(x[0]) \\
&u_1[2] = f_2(x[0], v_2[1]) \\
&\vdots \\
&u_1[k-1] = f_{k-1}(x[0], v_2[1:k-2])
\end{aligned}
$$

The receiver's observations $y_2[n]$ are given as follows.

$$
\begin{aligned}
&y_2[0] = x[0] + v_2[0] \\
&y_2[1] = ax[0] + u_1[0] + v_2[1] \\
&y_2[2] = a^2 x[0] + au_1[0] + u_1[1] + v_2[2] \\
&\vdots \\
&y_2[k-1] = a^{k-1}x[0] + a^{k-2}u_1[0] + \cdots + u_1[k-2] + v_2[k-1]
\end{aligned}
$$

The receiver generates an estimate $\widehat{x[0]}$ of the state $x[0]$ based on its received signal $y_2[1:k-1]$, i.e. $\widehat{x[0]} = g(y_2[1:k-1])$. The objective of the system is minimizing the quadratic estimation error, $\mathbb{E}[(x[0] - \widehat{x[0]})^2]$.

This problem can be more compactly represented as Figure 4.22 by thinking of the MIMO channel as a channel with memory. As shown in [22], feedback only increases the capacity at most a half bit per time step. However, in this problem we are using the channels for k time steps, so we still have to justify that the feedback does not increase the capacity too much. The following lemma explicitly computes an information-theoretic cutset bound for this communication problem and gives a reasonable bound on the rate-distortion tradeoff.

Lemma 4.11. *Let's consider Problem I of Figure 4.21.*
(i) Let $x[0] \sim \mathcal{N}(0, \sigma_0^2)$ and $v_2[n] \sim \mathcal{N}(0, \sigma_v^2)$. Let $w \in \mathbb{R}$ satisfy $|\frac{1}{a^2 w}| < 1$ and the input power constraint is

$$(1-w)\mathbb{E}[u_1^2[0]] + (1-w)w\mathbb{E}[u_1^2[1]] + \cdots + (1-w)w^{k-2}\mathbb{E}[u_1^2[k-2]] \leq P$$

Then, the estimation error of $x[0]$ based on $y_2[0:k-1]$ is lower bounded by

$$\mathbb{E}[(x[0] - \widehat{x[0]})^2] \geq \frac{\sigma_0^2}{2^{2I_k}}$$

where

$$I_k = \frac{k}{2}\log(1+\frac{1}{k\sigma_v^2}(\frac{2a^{2(k-1)}\sigma_0^2}{1-a^{-2}} + \frac{2a^{k-2}}{1-a^{-2}}\frac{P}{(1-\frac{1}{a^2 w})(1-w)})).$$

(ii) Consider the same problem as (i) except that $v_2[k-1] \sim \mathcal{N}(0, \sigma_v'^2)$, i.e. only the last observation noise variance is different. Then, the estimation error based on $y_2[0:k-1]$ is lower bounded by

$$\mathbb{E}[(x[0]-\widehat{x[0]})^2] \geq \frac{\sigma_0^2}{2^{2I_k'}}$$

where

$$I_k' = I_{k-1} + \frac{1}{2}\log(1+\frac{1}{\sigma_v'^2}(2a^{2(k-1)}\sigma_0^2 + 2\frac{a^{2(k-2)}}{1-\frac{1}{a^2 w}}\frac{P}{1-w})).$$

(iii) Consider the same problem as (ii) except that $v_2[k-1] \sim Unif[-\sigma_v', \sigma_v']$, i.e. the last observation is a uniform random variable. Then, the estimation error based on $y_2[0:k-1]$ is lower bounded by

$$\mathbb{E}[(x[0]-\widehat{x[0]})^2] \geq \frac{\sigma_0^2}{2^{2I_k''}}$$

where

$$I_k'' = I_k' + \frac{1}{2}\log(\frac{\pi e}{2}).$$

Proof. (i) First, we can lower bound the estimation error as follows:

$$\begin{aligned}
&\frac{1}{2}\log(2\pi e\mathbb{E}[(x[0]-\widehat{x[0]})^2]) \\
&\geq h(x[0]-\widehat{x[0]}|y_2[0:k-1]) \\
&= h(x[0]|y_2[0:k-1]) \\
&= h(x[0]) - I(x[0];y_2[0:k-1]) \\
&\geq \frac{1}{2}\log(2\pi e\sigma_0^2) - I(x[0];y_2[0:k-1]). \qquad (4.15)
\end{aligned}$$

We will upper bound the mutual information. Let's first upper bound the received signal power. Since $u_1[n]$ is a strictly causal function of $v_2[n]$,

$$\mathbb{E}[y_2^2[n]] \leq 2\mathbb{E}[(a^n x[0])^2] + 2\mathbb{E}[(a^{k-2}u_1[0] + u_1[n-1])^2] + \mathbb{E}[v_2^2[n]].$$

By Lemma 4.10, we have

$$\begin{aligned}
&\mathbb{E}[(a^{n-1}u_1[0] + \cdots + u_1[n-1])^2] \\
&\leq \frac{a^{2(n-1)}}{1-\frac{1}{a^2 w}}\frac{P}{1-w}.
\end{aligned}$$

Therefore, the received signal power is upper bounded as

$$\mathbb{E}[y_2^2[n]] \leq 2a^{2n}\sigma_0^2 + 2\frac{a^{2(n-1)}}{1-\frac{1}{a^2w}}\frac{P}{1-w} + \mathbb{E}[v_2^2[n]].$$

Thus, we can conclude

$$\begin{aligned}
&\sum_{0\leq n\leq k-1} \mathbb{E}[y_2^2[n]] \\
&\leq \sum_{0\leq n\leq k-1} 2a^{2n}\sigma_0^2 + 2\frac{a^{2(n-1)}}{1-\frac{1}{a^2w}}\frac{P}{1-w} + \sigma_v^2 \\
&= 2(1+\cdots+a^{2(k-1)})\sigma_0^2 + \sum_{0\leq n\leq k-1} 2\frac{a^{2(n-1)}}{1-\frac{1}{a^2w}}\frac{P}{1-w} + k\sigma_v^2 \\
&= 2a^{2(k-1)}\frac{1-a^{-2k}}{1-a^{-2}}\sigma_0^2 + 2a^{k-2}\frac{1-a^{-2k}}{1-a^{-2}}\frac{P}{(1-\frac{1}{a^2w})(1-w)} + k\sigma_v^2 \\
&\leq \frac{2a^{2(k-1)}\sigma_0^2}{1-a^{-2}} + \frac{2a^{k-2}}{1-a^{-2}}\frac{P}{(1-\frac{1}{a^2w})(1-w)} + k\sigma_v^2.
\end{aligned}$$

Using this, we can upper bound the mutual information.

$$\begin{aligned}
&I(x[0];y_2[0:k-1]) \\
&\leq h(y_2[0:k-1]) - h(y_2[0:k-1]|x[0]) \\
&\leq \sum_{0\leq n\leq k-1} h(y_2[n]) - \sum_{0\leq n\leq k-1} h(v_2[n]) \\
&\leq \sum_{0\leq n\leq k-1} \frac{1}{2}\log(2\pi e\mathbb{E}[y_2[n]^2]) - \frac{k-1}{2}\log(2\pi e\sigma_v^2) \\
&= \frac{1}{2}\log(\prod_{0\leq n\leq k-1}\frac{\mathbb{E}[y_2[n]^2]}{\sigma_v^2}) \\
&\leq \frac{1}{2}\log\left((\frac{1}{k}\sum_{0\leq n\leq k-1}\frac{\mathbb{E}[y_2[n]^2]}{\sigma_v^2})^{k-1}\right) \\
&(\because \text{geometric mean and arithmetic mean}) \\
&\leq \frac{1}{2}\log\left((1+\frac{1}{k\sigma_v^2}(\frac{2a^{2(k-1)}\sigma_0^2}{1-a^{-2}} + \frac{2a^{k-2}}{1-a^{-2}}\frac{P}{(1-\frac{1}{a^2w})(1-w)}))^k\right)
\end{aligned} \tag{4.16}$$

The last term is I_k. By plugging (4.16) into (4.15), we get

$$\mathbb{E}[(x[0]-\widehat{x[0]})^2] \geq \frac{\sigma_0^2}{2^{2I_k}}$$

which finishes the proof.

(ii) We have

$$
\begin{aligned}
&I(x[0]; y_2[0:k-1]) \\
&\leq h(y_2[0:k-1]) - h(y_2[0:k-1]|x[0]) \\
&\leq \sum_{0\leq n\leq k-1} h(y_2[n]) - \sum_{0\leq n\leq k-1} h(v_2[n]) \\
&\leq \sum_{0\leq n\leq k-1} \frac{1}{2}\log(2\pi e\mathbb{E}[y_2[n]^2]) - \frac{k-1}{2}\log(2\pi e\sigma_v^2) - \frac{1}{2}\log(2\pi e\sigma_v'^2) \\
&= \frac{1}{2}\log(\prod_{0\leq n\leq k-2} \frac{\mathbb{E}[y_2[n]^2]}{\sigma_v^2}) + \frac{1}{2}\log(\frac{\mathbb{E}[y_2[k-1]^2]}{\sigma_v'^2}) \\
&\leq \frac{1}{2}\log(\frac{1}{k-1}\sum_{0\leq n\leq k-2} \frac{\mathbb{E}[y_2[n]^2]}{\sigma_v^2})^{k-1} + \frac{1}{2}\log(\frac{\mathbb{E}[y_2[k-1]^2]}{\sigma_v'^2}) \\
&(\because \text{geometric mean and arithmetic mean}) \\
&\leq \frac{1}{2}\log(1 + \frac{1}{(k-1)\sigma_v^2}(\frac{2a^{2(k-2)}\sigma_0^2}{1-a^{-2}} + \frac{2a^{k-3}}{1-a^{-2}}\frac{P}{(1-\frac{1}{a^2w})(1-w)}))^{k-1} \\
&+ \frac{1}{2}\log(1 + \frac{1}{\sigma_v'^2}(2a^{2(k-1)}\sigma_0^2 + 2\frac{a^{2(k-2)}}{1-\frac{1}{a^2w}}\frac{P}{1-w})).
\end{aligned} \quad (4.17)
$$

The last term is I'_k. By plugging (4.17) into (4.15), we get

$$\mathbb{E}[(x[0] - \widehat{x[0]})^2] \geq \frac{\sigma_0^2}{2^{2I'_k}}$$

which finishes the proof.

(iii) We can repeat the proof of (ii) replacing the distribution of $v_2[k-1]$ by uniform. □

In this lemma, the bound of (ii) is tighter than that of (i) since it excludes the last observation in the arithmetic-geometric inequality, but it is harder to compute. We also allow the variance of the last observation noise to be different from the other ones, since we will replace it with another distribution to adapt large deviation ideas.[16]

4.6.5 Lower bound on the optimal cost based on Witsenhausen's counterexample

Now, we can combine the previous results to derive a lower bound that will approximately match with s-stage signaling strategies. We will derive a lower bound on the weighted average cost of Problem B, i.e. we will find functions $D_{L,i}(\widetilde{P_1}, \widetilde{P_2})$ such that

$$\inf_{u_1,u_2} \limsup_{N\to\infty} \frac{1}{N}\sum_{0\leq n<N} q\mathbb{E}[x^2[n]] + r_1\mathbb{E}[u_1^2[n]] + r_2\mathbb{E}[u_2^2[n]] \geq \min_{\widetilde{P_1},\widetilde{P_2}\geq 0} qD_{L,i}(\widetilde{P_1}, \widetilde{P_2}) + r_1\widetilde{P_1} + r_2\widetilde{P_2}.$$

[16] Even though large deviation ideas usually introduce a sequence of atypical noise, here the SNR of the last observation dominates the SNR of all the other observations. Thus, it is enough to introduce atypically large noise only to the last observation.

Here, the lower bounds $D_{L,i}(\widetilde{P_1}, \widetilde{P_2})$ can be thought as a lower bound on $D(P_1, P_2)$, the power-disturbance tradeoff, of Problem F. The first bound $D_{L,1}$ is given in the following lemma, and the rest will be given in Lemma 4.13 of page 208.

Lemma 4.12. *Define $S_{L,1}$ as the set of $(k_1, k_2, k, \sigma'_{v2}, \alpha, \Sigma)$ such that*

$$\begin{aligned}
&k_1, k_2, k \in \mathbb{N}, \sigma'_{v2}, \alpha, \Sigma \in \mathbb{R}_+, \\
&k_1 \geq 1, k_2 - k_1 - 1 \geq 0, k \geq k_2, \\
&\sigma'_{v2} \geq 0, 0 \leq \alpha \leq 1, \\
&0 \leq \Sigma \leq \begin{cases} 1 & \text{when } k_1 = 1 \\ \frac{a^{2(k_1-1)}\sigma_{v1}^2}{(1+\frac{\sigma_{v1}^2}{\sigma_{v2}^2})(\frac{a^{2(k_1-2)}(1-a^{-2(k_1-1)})}{1-a^{-2}})+\sigma_{v1}^2} & \text{when } k_1 \geq 2 \end{cases}
\end{aligned}$$

We also define $D_{L,1}(\widetilde{P_1}, \widetilde{P_2}; k_1, k_2, k, \sigma'_{v2}, \alpha, \Sigma)$ as follows:

$$\begin{aligned}
&D_{L,1}(\widetilde{P_1}, \widetilde{P_2}; k_1, k_2, k, \sigma'_{v2}, \alpha, \Sigma) \\
&:= \alpha\Big(\sqrt{c\frac{a^{2(k-k_1)}\Sigma}{2^{2I'(\widetilde{P_1})}}} - \sqrt{c\frac{a^{2(k-k_1-1)}(1-(2.5a^{-2})^{k_2-k_1})}{1-2.5a^{-2}}\frac{\widetilde{P_1}}{1-2.5^{-1}}} \\
&- \sqrt{\frac{a^{2(k-k_2-1)}(1-(2.5a^{-2})^{k-k_2})}{1-2.5a^{-2}}\frac{2.5^{k_2-k_1}\widetilde{P_1}}{1-2.5^{-1}}} - \sqrt{\frac{a^{2(k-k_2-1)}(1-(2.5a^{-2})^{k-k_2})}{1-2.5a^{-2}}\frac{\widetilde{P_2}}{1-2.5^{-1}}}\Big)_+^2 \\
&+ (1-\alpha)\Big(\sqrt{\frac{a^{2(k-k_1-1)}\Sigma}{2^{2I''(\widetilde{P_1})}}} - \sqrt{\frac{a^{2(k-k_1-2)}(1-(2.5a^{-2})^{k-k_1-1})}{1-2.5a^{-2}}\frac{2.5\widetilde{P_1}}{1-2.5^{-1}}} \\
&- \sqrt{\frac{a^{2(k-k_2-1)}(1-(2.5a^{-2})^{k-k_2})}{1-2.5a^{-2}}\frac{\widetilde{P_2}}{1-2.5^{-1}}}\Big)_+^2 + 1
\end{aligned} \tag{4.18}$$

where

$$I''(\widetilde{P_1}) = \begin{cases} \frac{k_2-k_1-1}{2}\log(1+\frac{1}{(k_2-k_1-1)\sigma_{v2}^2}(\frac{2a^{2(k_2-2-k_1)}}{1-a^{-2}}\Sigma + \frac{2a^{2(k_2-3-k_1)}}{1-a^{-2}}\frac{2.5\widetilde{P_1}}{(1-2.5a^{-2})(1-2.5^{-1})})) \\ \quad \text{if } k_2-k_1-1>0 \\ 0 \quad \text{if } k_2-k_1-1=0 \end{cases}$$

$$I'(\widetilde{P_1}) = I''(\widetilde{P_1}) + \frac{1}{2}\log(1+\frac{1}{\sigma'^2_{v2}}(2a^{2(k_2-1-k_1)}\Sigma + 2\frac{a^{2(k_2-2-k_1)}\widetilde{P_1}}{(1-2.5a^{-2})(1-2.5^{-1})})) + \frac{1}{2}\log(\frac{2\pi e}{4})\mathbf{1}(\sigma_{v2} \neq \sigma'_{v2})$$

$$c = \begin{cases} \frac{2\sigma'_{v2}}{\sqrt{2\pi}\sigma_{v2}}\exp(-\frac{\sigma'^2_{v2}}{2\sigma_{v2}^2}) & \text{if } \sigma_{v2} \neq \sigma'_{v2} \\ 1 & \text{if } \sigma_{v2} = \sigma'_{v2} \end{cases}$$

Let $|a| \geq 2.5$. Then, for all $q, r_1, r_2 \geq 0$, the minimum cost (4.1) of Problem B is lower bounded as follows:

$$\begin{aligned}
&\inf_{u_1,u_2}\limsup_{N\to\infty}\frac{1}{N}\sum_{0\leq n<N} q\mathbb{E}[x^2[n]] + r_1\mathbb{E}[u_1^2[n]] + r_2\mathbb{E}[u_2^2[n]] \\
&\geq \sup_{(k_1,k_2,k,\sigma'_{v2},\alpha,\Sigma)\in S_{L,1}} \min_{\widetilde{P_1},\widetilde{P_2}\geq 0} qD_{L,1}(\widetilde{P_1}, \widetilde{P_2}; k_1, k_2, k, \sigma'_{v2}, \alpha, \Sigma) + r_1\widetilde{P_1} + r_2\widetilde{P_2}.
\end{aligned}$$

Proof. For simplicity, we assume $a \geq 2.5$, $k_1 \geq 2$, $k_2 - k_1 - 1 > 0$, $k > k_2$, $\sigma_{v2} \neq \sigma'_{v2}$. The remaining cases when $a \leq -2.5$ or $k_1 = 1$ or $k_2 - k_1 - 1 = 0$ or $k = k_2$ or $\sigma_{v2} = \sigma'_{v2}$ easily follow with minor modifications.

• Geometric Slicing: We first apply the geometric slicing idea of Section 4.6.1 to get a finite-horizon problem. By setting $\alpha = 2.5^{-1}$ in Lemma 4.8, the average cost is lower bounded by

$$\inf_{u_1,u_2} (q\mathbb{E}[x^2[k]]$$
$$+ r_1 \underbrace{((1-2.5^{-1})\mathbb{E}[u_1^2[k_1]] + (1-2.5^{-1})2.5^{-1}\mathbb{E}[u_1^2[k_1+1]] + \cdots + (1-2.5^{-1})2.5^{-k+1+k_1}\mathbb{E}[u_1^2[k-1]])}_{:=\widetilde{P_1}}$$

$$+ r_2 \underbrace{((1-2.5^{-1})\mathbb{E}[u_2^2[k_2]] + (1-2.5^{-1})2.5^{-1}\mathbb{E}[u_2^2[k_2+1]] + \cdots + (1-2.5^{-1})2.5^{-k+1+k_2}\mathbb{E}[u_2^2[k-1]])}_{:=\widetilde{P_2}})$$

Here, we denote the second and the third terms as $\widetilde{P_1}$ and $\widetilde{P_2}$ respectively. As we mentioned in Figure 4.17, 4.18 and 4.19, we will relax the problem in two different ways — one with state disturbance $w[0]$ and the other one with $w[1]$. Let's start with the former.

• Large deviation idea: As mentioned in Section 4.6.3, we will apply large deviation ideas[17] to $v_2[k_2-1]$. For this, we write $v_2[k_2-1]$ as a mixture of two independent random variables:

$$v_2[k_2-1] = C \cdot v_2'[k_2-1] + (1-C)v_2''[k_2-1]$$

where $C, v_2'[k_2-1], v_2''[k_2-1]$ are independent random variables whose distributions are given as follows:

$$v_2'[k_2-1] \sim Unif[-\sigma'_{v2}, \sigma'_{v2}]$$
$$f_{v_2''[k_2-1]}(v) = \begin{cases} \frac{1}{1-c}\frac{1}{\sqrt{2\pi}\sigma_{v2}}\exp(-\frac{v^2}{2\sigma_{v2}^2}) & \text{for } |v| > \sigma'_{v2} \\ \frac{1}{1-c}\frac{1}{\sqrt{2\pi}\sigma_{v2}}(\exp(-\frac{v^2}{2\sigma_{v2}^2}) - \exp(-\frac{\sigma'^2_{v2}}{2\sigma_{v2}^2})) & \text{for } |v| \leq \sigma'_{v2} \end{cases}$$
$$C = \begin{cases} 1 & \text{w.p. } c \\ 0 & \text{w.p. } 1-c \end{cases}$$

where $c = \frac{2\sigma'_{v2}}{\sqrt{2\pi}\sigma_{v2}}\exp(-\frac{\sigma'^2_{v2}}{2\sigma_{v2}^2})$.

• Three stage division: As mentioned in Section 4.6.2, we will divide the finite-horizon problem into three time intervals. The following definitions of U_{ij} correspond to the first and second

[17] As mentioned before, large deviation theory usually replaces the whole noise sequence with a "typically atypical" one. However, for simplicity of computation, we will only replace the last observation noise. The Gaussian observation noise $v_2[k_2-1]$ will behave like a uniform observation noise with larger variance with a certain probability. Thus, we can replace $v_2[k_2-1]$ with a uniform random variable with larger variance by multiplying by the corresponding probability. See [37] for the details of the idea.

controller's input in these three intervals shown in Figure 4.16, where the indices i and j represent the controllers and time intervals respectively.

$$\begin{aligned}
W &:= aw[k-2] + \cdots + a^{k-1}w[0] \\
U_{11} &:= a^{k-2}u_1[1] + \cdots + a^{k-k_1}u_1[k_1-1] \\
U_{12} &:= a^{k-k_1-1}u_1[k_1] + \cdots + a^{k-k_2}u_1[k_2-1] \\
U_{13} &:= a^{k-k_2-1}u_1[k_2] + \cdots + u_1[k-1] \\
U_{21} &:= a^{k-2}u_2[1] + \cdots + a^{k-k_1}u_2[k_1-1] \\
U_{22} &:= a^{k-k_1-1}u_2[k_1] + \cdots + a^{k-k_2}u_2[k_2-1] \\
U_{23} &:= a^{k-k_2-1}u_2[k_2] + \cdots + u_2[k-1] \\
\overline{W} &:= (w[k-1], w[k-2], \cdots, w[1])
\end{aligned}$$

The goal in this proof is grouping control inputs into U_{ij}, where each U_{ij} can be thought as either power-limited or information-limited inputs. By expanding $x[n]$, we reveal the effects of the controller inputs on the state, and then isolate (and bound) their effects according to their characteristics.

• Power-Limited Interval: Let's first handle the third interval using Cauchy-Schwarz inequalities. Notice that

$$\begin{aligned}
x[k] = w[k-1] &+ aw[k-2] + \cdots + a^{k-1}w[0] \\
&+ u_1[k-1] + au_1[k-2] + \cdots + a^{k-2}u_1[1] \\
&+ u_2[k-1] + au_2[k-2] + \cdots + a^{k-2}u_2[1]
\end{aligned}$$

Therefore, by Lemma 4.1

$$\begin{aligned}
\mathbb{E}[x^2[k]] &= \mathbb{E}[(W + U_{11} + U_{12} + U_{13} + U_{21} + U_{22} + U_{23})^2] + \mathbf{E}[w^2[k-1]] \\
&\geq (\sqrt{\mathbb{E}[(W + U_{11} + U_{12} + U_{21} + U_{22})^2]} - \sqrt{\mathbb{E}[U_{13}^2]} - \sqrt{\mathbb{E}[U_{23}^2]})_+^2 + 1 \qquad (4.19)
\end{aligned}$$

Here, we can notice that $\mathbb{E}[(W + U_{11} + U_{12} + U_{21} + U_{22})^2]$ is not affected by the controllers' inputs in the third interval.

• First controller's input in Witsenhausen's interval: We will also separate out the effect of the power-limited (first controller's) input in the second interval, U_{12}, and introduce large deviation ideas.

$$\mathbb{E}[(W+U_{11}+U_{12}+U_{21}+U_{22})^2]$$
$$=\mathbb{E}[\mathbb{E}[(W+U_{11}+U_{12}+U_{21}+U_{22})^2|C]]$$
$$=\mathbb{E}[(W+U_{11}+U_{12}+U_{21}+U_{22})^2|C=1]\mathbb{P}(C=1)+\mathbb{E}[(W+U_{11}+U_{12}+U_{21}+U_{22})^2|C=0]\mathbb{P}(C=0)$$
$$\geq \mathbb{E}[(W+U_{11}+U_{12}+U_{21}+U_{22})^2|C=1]\mathbb{P}(C=1)$$
$$=c\cdot\mathbb{E}[(W+U_{11}+U_{12}+U_{21}+U_{22})^2|C=1]$$
$$\geq c(\sqrt{\mathbb{E}[(W+U_{11}+U_{21}+U_{22})^2|C=1]}-\sqrt{\mathbb{E}[U_{12}^2|C=1]})_+^2 \tag{4.20}$$

Here, we can notice that by the causality of the system, C only affects the inputs $u_2[k_2-1]$ and $u_1[k_2]$. Thus, $u_2[1:k_2-2]$ and $u_1[1:k_2-1]$ are independent of C. We can also notice $\mathbb{E}[(W+U_{11}+U_{21}+U_{22})^2|C=1]$ has only information-limited inputs.

• Information-Limited Interval: Using Lemma 4.9, we will bound the remaining uncertainty of the state after the information-limited interval. Since we will grant all disturbances except $w[0]$ as side-information, we denote the relevant observations as $y_1'[n]$ and $y_2'[n]$. Formally, let $y_1'[n]$, $y_2'[n]$, W', W'', U_{22}', U_{22}'' be as follows:

$$y_1'[n]:=a^{n-1}w[0]+v_1[n]$$
$$y_2'[n]:=a^{n-1}w[0]+v_2[n]$$
$$W':=W-\mathbb{E}[W|y_1'[1:k_1-1],y_2'[1:k_1-1],\overline{W},C=1]$$
$$W'':=\mathbb{E}[W|y_1'[1:k_1-1],y_2'[1:k_1-1],\overline{W},C=1]$$
$$U_{22}':=U_{22}-\mathbb{E}[U_{22}|y_1'[1:k_1-1],y_2'[1:k_1-1],\overline{W},C=1]$$
$$U_{22}'':=\mathbb{E}[U_{22}|y_1'[1:k_1-1],y_2'[1:k_1-1],\overline{W},C=1]$$

Here we can notice $W,y_1'[1:k_1-1],y_2'[1:k_1-1],\overline{W}$ are independent of C and

$$W'=a^{k-1}w[0]-\mathbb{E}[a^{k-1}w[0]|y_1'[1:k_1-1],y_2'[1:k_1-1]].$$

Since $w[0],y_1'[1:k_1-1],y_2'[1:k_1-1],\overline{W}$ are jointly Gaussian, W' is independent from $y_1'[1:k_1-1],y_2'[1:k_1-1],\overline{W}$. By Lemma 4.9 we have

$$\mathbb{E}[W'^2|C=1]$$
$$=\mathbb{E}[(a^{k-1}w[0]-\mathbb{E}[a^{k-1}w[0]|y_1'[1:k_1-1],y_2'[1:k_1-1]])^2]$$
$$=\frac{a^{2(k-1)}\sigma_{v1}^2}{(1+\frac{\sigma_{v1}^2}{\sigma_{v2}^2})(\frac{a^{2(k_1-2)}(1-a^{-2(k_1-1)})}{1-a^{-2}})+\sigma_{v1}^2} \tag{4.21}$$

This lower bounds the uncertainty in the state due to $w[0]$ after the state has been observed through $y_1'[1:k_1-1]$ and $y_2'[1:k_1-1]$.

Note that $y_1[1:k_1-1], y_2[1:k_1-1], \overline{W}$ are functions of $y'_1[1:k_1-1], y'_2[1:k_1-1], \overline{W}$. Therefore, U_{11} and U_{21} are also functions of $y'_1[1:k_1-1], y'_2[1:k_1-1], \overline{W}$. Since (W', U'_{22}) are orthogonal to all functions of $(y'_1[1:k_1-1], y'_2[1:k_1-1], \overline{W})$, (W', U'_{22}) are also orthogonal to $(W'', U_{11}, U_{21}, U''_{22})$. Moreover, since $\mathbb{E}[W' + U'_{22}] = 0$ and the conditioning on $C = 1$ can be ignored due to causality, we can conclude

$$\begin{aligned}
&\mathbb{E}[(W + U_{11} + U_{21} + U_{22})^2 | C = 1] \\
&= \mathbb{E}[(W' + W'' + U_{11} + U_{21} + U'_{22} + U''_{22})^2 | C = 1] \\
&= \mathbb{E}[(W' + U'_{22})^2 | C = 1] + \mathbb{E}[(W'' + U_{11} + U_{21} + U''_{22})^2 | C = 1] \\
&\geq \mathbb{E}[(W' + U'_{22})^2 | C = 1].
\end{aligned}$$

In the last term, the effect of the information-limited interval inputs is separated out.

• Second controller's input in Witsenhausen's interval: We will bound the remaining uncertainty of the state after it has been estimated by the second controller in the second time interval. For this, we will reduce the problem to the state amplification problem of Section 4.6.4, and apply Lemma 4.11.

U'_{22} is a function of $y_2[1:k_2-1], y'_1[1:k_1-1], y'_2[1:k_1-1], \overline{W}$. Here, $y'_1[1:k_1-1], y'_2[1:k_1-1], \overline{W}$ are independent from W' and $y_2[1:k_1-1]$ is a function of $y'_1[1:k_1-1], y'_2[1:k_1-1], \overline{W}$. Therefore, only $y_2[k_1:k_2-1]$ are dependent on W'. Moreover, $y_1[1:k_1-1]$ —and therefore, $u_1[1:k_1-1]$— is a function of $y'_1[1:k_1-1], y'_2[1:k_1-1], \overline{W}$, so they are also independent from W'.

Now, we can subtract the independent part from W' from the observation $y_2[k_1, k_2-1]$ without losing information about the state. First, consider $y_2[k_1]$.

$$\begin{aligned}
&y_2[k_1] - (w[k_1-1] + aw[k_1-2] + \cdots + a^{k_1-2}w[1]) - \mathbb{E}[a^{k_1-1}w[0] | y'_1[1:k_1-1], y'_2[1:k_1-]] \\
&- (u_1[k_1-1] + au_1[k_1-2] + \cdots + a^{k_1-2}u_1[1]) \\
&- (u_1[k_1-1] + au_1[k_1-2] + \cdots + a^{k_1-2}u_1[1]) \\
&= a^{k_1-1}w[0] - \mathbb{E}[a^{k_1-1}w[0] | y'_1[1:k_1-1], y'_2[1:k_1-1]] + v_2[k_1] \\
&= a^{k_1-k}W' + v_2[k_1]
\end{aligned}$$

Likewise, we can subtract the independent (from W') part from $y_2[k_1+1]$. Furthermore, $u_2[k_1]$ can also be subtracted from $y_2[k_1+1]$ without losing information since the second controller already knows about $u_2[k_1]$. Thus, the information about W' in $y_2[k_1+1]$ is in

$$\begin{aligned}
&a^{k_1}w[0] - \mathbb{E}[a^{k_1}w[0] | y'_1[1:k_1-1], y'_2[1:k_1-1]] + u_1[k_1] + v_2[k_1+1] \\
&= a^{k_1-k+1}W' + u_1[k_1] + v_2[k_1+1].
\end{aligned}$$

In the same way, we can extract the relevant information about W' from the observations $y_2[n]$. It is worth to mention that conditioned on $C = 1$, $v_2[k_2-1]$ is replaced by $v'_2[k_2-1]$, and thus the information about W' in $y_2[k_2-1]$ is

$$\begin{aligned}
&a^{k_2-2}w[0] - \mathbb{E}[a^{k_2-2}w[0]|y_1'[1:k_1-1], y_2'[1:k_1-1]] \\
&+ u_1[k_2-1] + au_1[k_2-2] + \cdots + a^{k_2-k_1-1}u_1[k_1] + v_2'[k_2-1] \\
&= a^{k_2-k-1}W' + u_1[k_2-1] + au_1[k_2-2] + \cdots + a^{k_2-k_1-1}u_1[k_1] + v_2'[k_2-1].
\end{aligned}$$

Moreover, as we mentioned, the conditioning $C = 1$ does not affect $u_1[k_1 : k_2 - 1]$ by causality. We have

$$\begin{aligned}
&\mathbb{E}[(1-2.5^{-1})u_1^2[k_1] + (1-2.5^{-1})2.5^{-1}u_1^2[k_1+1] + \cdots + (1-2.5^{-1})2.5^{-k_2+k_1+1}u_1^2[k_2-1]|C=1] \\
&\leq \widetilde{P_1} \leq 2.5\widetilde{P_1}.
\end{aligned}$$

Therefore, we can see that after removing the independent (from W') part from $y_2[k_1 : k_2 - 1]$ the problem reduces to the state amplification problem of Section 4.6.4. By plugging $x[0] = a^{k_1-1}w[0]$, $\sigma_v = \sigma_{v2}$, $\sigma_v' = \sigma_{v2}'$, $k = k_2 - k_1$, $w = 2.5^{-1}$, $P = 2.5\widetilde{P_1}$ and $\sigma_0^2 = \Sigma$ (which comes from (4.21)) in Lemma 4.11 (iii), we have[18]

$$\mathbb{E}[(W' + U_{22}')^2|C=1] \geq \frac{a^{2(k-k_1)}\Sigma}{2^{2I'(\widetilde{P_1})}}. \tag{4.22}$$

• Power-Limited Inputs: As mentioned before, causality implies C is independent from $y_1[1 : k_2 - 1]$ and thus U_{12}. Then, we can upper bound the power of the power-limited inputs.

$$\begin{aligned}
&\mathbb{E}[U_{12}^2|C=1] = \mathbb{E}[U_{12}^2] \\
&= \mathbb{E}[(a^{k-k_1-1}u_1[k_1] + \cdots + a^{k-k_2}u_1[k_2-1])^2] \\
&= a^{2(k-k_2)}\mathbb{E}[(a^{k_2-k_1-1}u_1[k_1] + \cdots + u_1[k_2-1])^2] \\
&\leq \frac{a^{2(k-k_1-1)}(1-(2.5a^{-2})^{k_2-k_1})}{1-2.5a^{-2}}(\mathbb{E}[u_1^2[k_1]] + \cdots + 2.5^{-(k-k_1-1)}\mathbb{E}[u_1^2[k_2-1]]) \\
&\leq \frac{a^{2(k-k_1-1)}(1-(2.5a^{-2})^{k_2-k_1})}{1-2.5a^{-2}}\frac{\widetilde{P_1}}{1-2.5^{-1}}
\end{aligned} \tag{4.23}$$

where the first inequality comes from Lemma 4.10 with parameters $a = a$ and $b = 2.5^{-1}$. Likewise, by applying Lemma 4.10 with paramters $a = a$ and $b = 2.5^{-1}$, we have

$$\begin{aligned}
&\mathbb{E}[U_{13}^2] = \mathbb{E}[(a^{k-k_1-1}u_1[k_1] + \cdots + u_1[k-1])^2] \\
&\leq \frac{a^{2(k-k_2-1)}(1-(2.5a^{-2})^{k-k_2})}{1-2.5a^{-2}}(\mathbb{E}[u_1^2[k_2]] + \cdots + 2.5^{-(k-k_2-1)}\mathbb{E}[u_1^2[k-1]]) \\
&\leq \frac{a^{2(k-k_2-1)}(1-(2.5a^{-2})^{k-k_2})}{1-2.5a^{-2}}\frac{2.5^{k_2-k_1}\widetilde{P_1}}{1-2.5^{-1}}
\end{aligned} \tag{4.24}$$

[18] Here, we have to use (iii) of Lemma 4.11 instead of (i) since in the last observation the SNR (Signal-to-Noise ratio) is too big to apply an arithmetic-geometric inequality together with the previous observations.

and

$$\begin{aligned}\mathbb{E}[U_{23}^2] &= \mathbb{E}[(a^{k-k_2-1}u_2[k_2]+\cdots+u_2[k-1])^2]\\ &\le \frac{a^{2(k-k_2-1)}(1-(2.5a^{-2})^{k-k_2})}{1-2.5a^{-2}}(\mathbb{E}[u_2^2[k_2]]+\cdots+2.5^{-(k-k_2-1)}\mathbb{E}[u_2^2[k-1]])\\ &\le \frac{a^{2(k-k_2-1)}(1-(2.5a^{-2})^{k-k_2})}{1-2.5a^{-2}}\frac{\widetilde{P_2}}{1-2.5^{-1}}. \end{aligned} \tag{4.25}$$

• Lower bound from $w[0]$: Finally, by plugging (4.20), (4.22), (4.23), (4.24), (4.25) into (4.19)

$$\begin{aligned}\mathbb{E}[x^2[k]] \ge (&\sqrt{c\frac{a^{2(k-k_1)}\Sigma}{2^{2I'(\widetilde{P_1})}}} - \sqrt{c\frac{a^{2(k-k_1-1)}(1-(2.5a^{-2})^{k_2-k_1})}{1-2.5a^{-2}}\frac{\widetilde{P_1}}{1-2.5^{-1}}}\\ &-\sqrt{\frac{a^{2(k-k_2-1)}(1-(2.5a^{-2})^{k-k_2})}{1-2.5a^{-2}}\frac{2.5^{k_2-k_1}\widetilde{P_1}}{1-2.5^{-1}}}\\ &-\sqrt{\frac{a^{2(k-k_2-1)}(1-(2.5a^{-2})^{k-k_2})}{1-2.5a^{-2}}\frac{\widetilde{P_2}}{1-2.5^{-1}}})_+^2+1 \end{aligned} \tag{4.26}$$

• Lower bound from $w[1]$: As we mentioned in Figure 4.17, 4.18 and 4.19, we will repeat the above derivation for $w[1]$ instead of $w[0]$.

Let's denote

$$\begin{aligned}\widetilde{U}_{11} &:= a^{k-2}u_1[1]+\cdots+a^{k-k_1-1}u_1[k_1]\\ \widetilde{U}_{12} &:= a^{k-k_1-2}u_1[k_1+1]+\cdots+u_1[k-1]\\ \widetilde{U}_{21} &:= a^{k-2}u_2[1]+\cdots+a^{k-k_1-1}u_2[k_1]\\ \widetilde{U}_{22} &:= a^{k-k_1-2}u_2[k_1+1]+\cdots+a^{k-k_2}u_2[k_2-1]\\ \widetilde{W} &:= (w[k-1],w[k-2],\cdots,w[2],w[0])\end{aligned}$$

Compared with the previous case, $\widetilde{U}_{11}$ and $\widetilde{U}_{21}$ include extra input signals $u_1[k_1]$ and $u_2[k_1]$ since $w[1]$ is generated one time-step later than $w[0]$. $\widetilde{U}_{12}$ includes all power-limited inputs of the first controller.

Like before, $\widetilde{U}_{ij}$ groups the controller inputs into either information-limited or power-limited ones. Then, we will isolate the effect of the inputs $\widetilde{U}_{ij}$ to the state $x[n]$ according to their categories.

• Power-Limited Inputs: Like the previous case, we first isolate the power-limited inputs. However, unlike the previous case, we do not need to introduce any large deviation ideas. By Lemma 4.1,

$$\begin{aligned}\mathbb{E}[x^2[k]] &= \mathbb{E}[(W+\widetilde{U}_{11}+\widetilde{U}_{12}+\widetilde{U}_{21}+\widetilde{U}_{22}+U_{23})^2]+1\\ &\ge (\sqrt{\mathbb{E}[(W+\widetilde{U}_{11}+\widetilde{U}_{21}+\widetilde{U}_{22})^2]}-\sqrt{\mathbb{E}[\widetilde{U}_{12}^2]}-\sqrt{\mathbb{E}[U_{23}^2]})_+^2+1\end{aligned} \tag{4.27}$$

Now, the resulting $\mathbb{E}[(W+\widetilde{U}_{11}+\widetilde{U}_{21}+\widetilde{U}_{22})^2]$ has only information-limited inputs.

• Information-Limited Interval: Like before, we will bound the remaining uncertainty of the state after the information-limited interval using Lemma 4.9. Denote $\widetilde{y}_1[n]$ and $\widetilde{y}_2[n]$ as follows:

$$\begin{aligned}
&\widetilde{y}_1[1] := v_1[1] \\
&\widetilde{y}_2[1] := v_2[1] \\
&\text{For } n \geq 2 \\
&\widetilde{y}_1[n] := a^{n-2}w[1] + v_1[n] \\
&\widetilde{y}_2[n] := a^{n-2}w[1] + v_2[n] \\
&W_1' := W - \mathbb{E}[W|\widetilde{y}_1[1:k_1], \widetilde{y}_2[1:k_1], \widetilde{W}] \\
&W_1'' := \mathbb{E}[W|\widetilde{y}_1[1:k_1], \widetilde{y}_2[1:k_1], \widetilde{W}] \\
&\widetilde{U}_{22}' = \widetilde{U} - \mathbb{E}[\widetilde{U}_{22}|\widetilde{y}_1[1:k_1], \widetilde{y}_2[1:k_1], \widetilde{W}] \\
&\widetilde{U}_{22}'' = \mathbb{E}[\widetilde{U}_{22}|\widetilde{y}_1[1:k_1], \widetilde{y}_2[1:k_1], \widetilde{W}]
\end{aligned}$$

Here we can notice

$$W_1' = a^{k-2}w[1] - \mathbb{E}[a^{k-2}w[1]|\widetilde{y}_1[1:k_1], \widetilde{y}_2[1:k_1]]$$

Since $w[1], \widetilde{y}_1[1:k_1], \widetilde{y}_2'[1:k_1], \widetilde{W}$ are jointly Gaussian, W_1' is independent from $\widetilde{y}_1[1:k_1], \widetilde{y}_2'[1:k_1], \widetilde{W}$. By Lemma 4.9 we have

$$\begin{aligned}
&\mathbb{E}[W_1'^2] \\
&= \mathbb{E}[(a^{k-2}w[1] - \mathbb{E}[a^{k-2}w[1]|\widetilde{y}_1[1:k_1], \widetilde{y}_2[1:k_1]])^2] \\
&= \frac{a^{2(k-2)}\sigma_{v2}^2}{(1+\frac{\sigma_{v1}^2}{\sigma_{v2}^2})(\frac{a^{2(k_1-2)}(1-a^{-2(k_1-1)})}{1-a^{-2}}) + \sigma_{v1}^2}
\end{aligned} \tag{4.28}$$

This lower bounds the remaining state disturbance due to $w[1]$ after it is observed by $\widetilde{y}_1[1:k_1]$ and $\widetilde{y}_2[1:k_1]$.

Note that $y_1[1:k_1], y_2[1:k_1], \widetilde{W}$ are functions of $\widetilde{y}_1[1:k_1], \widetilde{y}_2[1:k_1], \widetilde{W}$. Therefore, $\widetilde{U}_{11}$ and $\widetilde{U}_{21}$ are also functions of $\widetilde{y}_1[1:k_1], \widetilde{y}_2[1:k_1], \widetilde{W}$. By repeating the previous argument, we can conclude

$$\begin{aligned}
&\mathbb{E}[(W + \widetilde{U}_{11} + \widetilde{U}_{21} + \widetilde{U}_{22})^2] \\
&= \mathbb{E}[(W_1' + W_1'' + \widetilde{U}_{11} + \widetilde{U}_{21} + \widetilde{U}_{22}' + \widetilde{U}_{22}'')^2] \\
&= \mathbb{E}[(W_1' + \widetilde{U}_{22}')^2] + \mathbb{E}[(W_1'' + \widetilde{U}_{11} + \widetilde{U}_{21} + \widetilde{U}_{22}'')^2] \\
&\geq \mathbb{E}[(W_1' + \widetilde{U}_{22}')^2].
\end{aligned}$$

In the last term, the effect of the information-limited inputs is separated out.

• Second controller's input in Witsenhausen's interval: Like before, we will reduce the problem to the state amplification problem of Section 4.6.4, and apply Lemma 4.11. Only the observations $y_2[k_1+1:k_2-1]$ are relevant to W_1'. Here, we also have the power constraint on u_1

$$\mathbb{E}[(1-2.5^{-1})u_1^2[k_1+1]+\cdots+2.5^{-k_2+k_1+2}u_1^2[k_2-1]] \leq 2.5\widetilde{P_1}$$

Like before, after removing the independent (from W_1') part from the observations $y_2[k_1+1:k_2-1]$, the problem reduces to the state amplification problem of Section 4.6.4. By plugging $x[0]=a^{k_1-1}w[1], \sigma_v=\sigma_{v2}, k=k_2-k_1-1, w=2.5^{-1}, P=2.5\widetilde{P_1}$ and $\sigma_0^2=\Sigma$ (which comes from (4.28)) to Lemma 4.11 (i), we have[19]

$$\mathbb{E}[(W_1'+\widetilde{U}_{22}')^2] \geq \frac{a^{2(k-k_1-1)}\Sigma}{2^{2I''(\widetilde{P_1})}}. \tag{4.29}$$

• Lower bound from $w[1]$: By applying Lemma 4.10 with the parameters $a=a$ and $b=2.5^{-1}$, we can upper bound the power of the power-limited inputs.

$$\begin{aligned}\mathbb{E}[\widetilde{U}_{12}^2] &= \mathbb{E}[(a^{k-k_1-2}u_1[k_1+1]+\cdots+u_1[k-1])^2]\\ &\leq \frac{a^{2(k-k_1-2)}(1-(2.5a^{-2})^{k-k_1-1})}{1-2.5a^{-2}}\frac{2.5\widetilde{P_1}}{1-2.5^{-1}}\end{aligned} \tag{4.30}$$

Therefore, by plugging (4.29), (4.30), (4.25) into (4.27) we get

$$\begin{aligned}\mathbb{E}[x^2[k]] \geq (&\sqrt{\frac{a^{2(k-k_1-1)}\Sigma}{2^{2I''(\widetilde{P_1})}}}-\sqrt{\frac{a^{2(k-k_1-2)}(1-(2.5a^{-2})^{k-k_1-1})}{1-2.5a^{-2}}\frac{2.5\widetilde{P_1}}{1-2.5^{-1}}}\\ &-\sqrt{\frac{a^{2(k-k_2-1)}(1-(2.5a^{-2})^{k-k_2})}{1-2.5a^{-2}}\frac{\widetilde{P_2}}{1-2.5^{-1}}})_+^2+1\end{aligned} \tag{4.31}$$

• Final Lower bound: By (4.26) and (4.31), for all $0 \leq \alpha \leq 1$

$$\begin{aligned}&\mathbb{E}[x^2[k]]\\ &\geq \alpha(\sqrt{c\frac{a^{2(k-k_1)}\Sigma}{2^{2I'(\widetilde{P_1})}}}-\sqrt{c\frac{a^{2(k-k_1-1)}(1-(2.5a^{-2})^{k_2-k_1})}{1-2.5a^{-2}}\frac{\widetilde{P_1}}{1-2.5^{-1}}}\\ &-\sqrt{\frac{a^{2(k-k_2-1)}(1-(2.5a^{-2})^{k-k_2})}{1-2.5a^{-2}}\frac{2.5^{k_2-k_1}\widetilde{P_1}}{1-2.5^{-1}}}-\sqrt{\frac{a^{2(k-k_2-1)}(1-(2.5a^{-2})^{k-k_2})}{1-2.5a^{-2}}\frac{\widetilde{P_2}}{1-2.5^{-1}}})_+^2\\ &+(1-\alpha)(\sqrt{\frac{a^{2(k-k_1-1)}\Sigma}{2^{2I''(\widetilde{P_1})}}}-\sqrt{\frac{a^{2(k-k_1-2)}(1-(2.5a^{-2})^{k-k_1-1})}{1-2.5a^{-2}}\frac{2.5\widetilde{P_1}}{1-2.5^{-1}}}\\ &-\sqrt{\frac{a^{2(k-k_2-1)}(1-(2.5a^{-2})^{k-k_2})}{1-2.5a^{-2}}\frac{\widetilde{P_2}}{1-2.5^{-1}}})_+^2+1\end{aligned}$$

□

In this lemma, the time-interval from 0 to k_1-1 corresponds to the information-limited interval in Figure 4.16. The time-interval from k_1 to k_2-1 corresponds to the Witsenhausen's interval in Figure 4.16. The time-interval from k_2 to k corresponds to the power-limited interval in Figure 4.16.

[19]Unlike the previous part, we apply (i) of Lemma 4.11 instead of (iii) since the SNR is small enough for all observations.

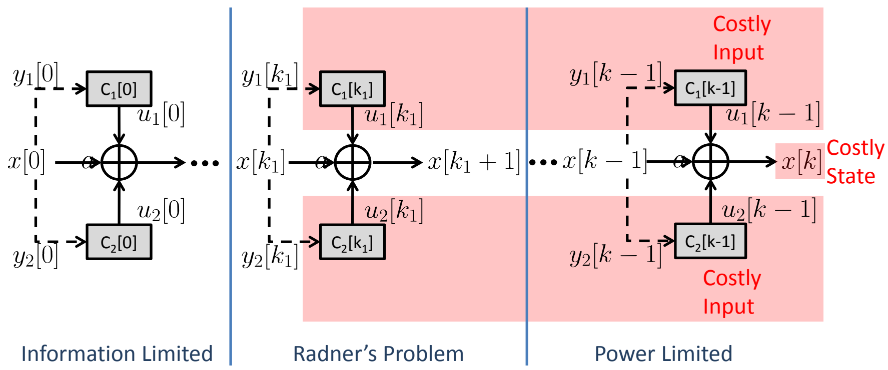


Figure 4.23: The general finite-horizon problem structure to justify the infeasibility of 0-stage signaling strategies. Like the one in Figure 4.16, the problem consists of three time intervals. However, unlike Figure 4.16, we can see Radner's problem between the information-limited and power-limited intervals.

4.6.6 Lower bound on the optimal cost based on Radner's problem

As we discussed in Section 4.3.1, Radner's problem cannot be understood using the binary deterministic models and thereby is fundamentally different from Witsenhausen's counterexample. Essentially, it says the communication between controllers requires at least one step delay, and for the observations obtained at the same time step, nonlinear strategies do not improve the performance. Therefore, so-called '0-stage signaling' is impossible.

Sine Radner's problem is a sub-block of the infinite-horizon problem 4.1, we also need a lower bound based on Radner's problem to bound the infinite-horizon problem within a constant ratio. Figure 4.23 shows the general structure of the lower bound for the case. As we discussed in Figure 4.16, the information-limited interval from time step 0 to k_1 is introduced due to the case $\sigma_{v1}^2 > 0$ and the power-limited interval from time step $k_1 + 1$ to k is introduced due to the case $P_2 < \infty$.

However, between these two time intervals, we can see the difference. Even though the first controller has better observations and the second has worse observations, if this significant unbalance between two controllers lasts for only one time step, implicit communication between the two controllers is nearly impossible and nonlinear strategies cannot help that much. To capture this effect, we replace the MIMO Witsenhausen's problem with Radner's problem.

Like Lemma 4.9, the following lemma gives a lower bound on the weighted average cost of Problem B when Witsenhausen's interval is replaced by Radner's.

Lemma 4.13. *Define a set $S_{L,2}$ as a set of (k_1, k, Σ) such that*

$$
\begin{aligned}
&k_1, k \in \mathbb{N}, \Sigma \in \mathbb{R}, \\
&k_1 \geq 1, k \geq k_1 + 1, \\
&0 \leq \Sigma \leq \begin{cases} 1 & k_1 = 1 \\ \frac{a^{2(k_1-1)}\sigma_{v1}^2}{(1+\frac{\sigma_{v1}^2}{\sigma_{v2}^2})(\frac{a^{2(k_1-2)}(1-a^{-2(k_1-1)})}{1-a^{-2}})+\sigma_{v1}^2} & k_1 \geq 2 \end{cases}
\end{aligned}
$$

We also define $D_{L,2}(\widetilde{P_1}, \widetilde{P_2}; k_1, k, \Sigma)$ as follows:

$$
\begin{aligned}
&D_{L,2}(\widetilde{P_1}, \widetilde{P_2}; k_1, k, \Sigma) \\
&:= \inf_{c_1, c_2 \in \mathbb{R}} (\sqrt{a^{2(k-k_1-1)}((a-c_1-c_2)^2\Sigma + c_1^2\sigma_{v1}^2 + c_2^2\sigma_{v2}^2)} \\
&- \sqrt{\frac{a^{2(k-k_1-2)}(1-(2.5a^{-2})^{k-k_1-1})}{1-2.5a^{-2}} \frac{\widetilde{P_1}}{(1-2.5^{-1})2.5^{-1}}} \\
&- \sqrt{\frac{a^{2(k-k_1-2)}(1-(2.5a^{-2})^{k-k_1-1})}{1-2.5a^{-2}} \frac{\widetilde{P_2}}{(1-2.5^{-1})2.5^{-1}}})_+^2 + 1 \\
&s.t.\ (1-2.5^{-1})c_1^2(\Sigma+\sigma_{v1}^2) \leq \widetilde{P_1} \\
&\quad (1-2.5^{-1})c_2^2(\Sigma+\sigma_{v2}^2) \leq \widetilde{P_2}
\end{aligned}
$$

Let $|a| \geq 2.5$. Then, for all $q, r_1, r_2 \geq 0$, the minimum cost (4.1) *of Problem B is lower bounded as follows:*

$$
\begin{aligned}
&\inf_{u_1,u_2} \limsup_{N\to\infty} \frac{1}{N} \sum_{0 \leq n < N} q\mathbb{E}[x^2[n]] + r_1\mathbb{E}[u_1^2[n]] + r_2\mathbb{E}[u_2^2[n]] \\
&\geq \sup_{(k_1,k,\Sigma)\in S_{L,2}} \min_{\widetilde{P_1},\widetilde{P_2}\geq 0} qD_{L,2}(\widetilde{P_1}, \widetilde{P_2}; k_1, k, \Sigma) + r_1\widetilde{P_1} + r_2\widetilde{P_2}.
\end{aligned}
$$

For $k_1, k \in \mathbb{N}$, define $S_{L,3}$, $S_{L,4}$, $D_{L,3}(\widetilde{P_1}, \widetilde{P_2}; k_1)$ and $D_{L,4}(\widetilde{P_1}, \widetilde{P_2}, k)$ as follows:

$$
\begin{aligned}
&S_{L,3} := \{k_1 \in \mathbb{N}\} \\
&S_{L,4} := \{k \in \mathbb{N} : k \geq 2\} \\
&D_{L,3}(\widetilde{P_1}, \widetilde{P_2}; k_1) := \max(\frac{a^{2(k_1-1)}\sigma_{v1}^2}{(1+\frac{\sigma_{v1}^2}{\sigma_{v2}^2})(\frac{a^{2(k_1-2)}(1-a^{-2(k_1-1)})}{1-a^{-2}})+\sigma_{v1}^2}, 1) \quad (4.32) \\
&D_{L,4}(\widetilde{P_1}, \widetilde{P_2}; k) := (\sqrt{a^{2(k-1)}} - \sqrt{\frac{a^{2(k-2)}}{1-2.5a^{-2}}\frac{\widetilde{P_1}}{1-2.5^{-1}}} - \sqrt{\frac{a^{2(k-2)}}{1-2.5a^{-2}}\frac{\widetilde{P_2}}{1-2.5^{-1}}})_+^2. \quad (4.33)
\end{aligned}
$$

Then, when $|a| \geq 2.5$, for all $q, r_1, r_2 \geq 0$, the minimum cost (4.1) *of Problem B is also lower bounded as follows:*

$$
\begin{aligned}
&\inf_{u_1,u_2} \limsup_{N\to\infty} \frac{1}{N} \sum_{0 \leq n < N} q\mathbb{E}[x^2[n]] + r_1\mathbb{E}[u_1^2[n]] + r_2\mathbb{E}[u_2^2[n]] \\
&\geq \sup_{k_1\in S_{L,3}} \min_{\widetilde{P_1},\widetilde{P_2}\geq 0} qD_{L,3}(\widetilde{P_1}, \widetilde{P_2}; k_1) + r_1\widetilde{P_1} + r_2\widetilde{P_2}
\end{aligned}
$$

and

$$\inf_{u_1,u_2} \limsup_{N\to\infty} \frac{1}{N} \sum_{0\leq n<N} q\mathbb{E}[x^2[n]] + r_1\mathbb{E}[u_1^2[n]] + r_2\mathbb{E}[u_2^2[n]]$$
$$\geq \sup_{k\in S_{L,4}} \min_{\widetilde{P_1},\widetilde{P_2}\geq 0} qD_{L,4}(\widetilde{P_1},\widetilde{P_2};k) + r_1\widetilde{P_1} + r_2\widetilde{P_2}.$$

Proof. For simplicity, we assume $a \geq 2.5$, $k_1 \geq 2$, $k > k_1 + 1$. The remaining cases when $a \leq -2.5$ or $k_1 = 1$ or $k = k_1 + 1$ easily follow with minor modifications.

• Geometric Slicing: We apply the geometric slicing idea of Section 4.6.1 to get a finite-horizon problem. By putting $\alpha = 2.5^{-1}$ and $k_2 = k_1$ to Lemma 4.8, the average cost is lower bounded by

$$\inf_{u_1,u_2} (q\mathbb{E}[x^2[k]]$$
$$+ r_1 \underbrace{((1-2.5^{-1})\mathbb{E}[u_1^2[k_1]] + (1-2.5^{-1})2.5^{-1}\mathbb{E}[u_1^2[k_1+1]] + \cdots + (1-2.5^{-1})2.5^{-k+1+k_1}\mathbb{E}[u_1^2[k-1]])}_{:=\widetilde{P_1}}$$
$$+ r_2 \underbrace{((1-2.5^{-1})\mathbb{E}[u_2^2[k_1]] + (1-2.5^{-1})2.5^{-1}\mathbb{E}[u_2^2[k_1+1]] + \cdots + (1-2.5^{-1})2.5^{-k+1+k_1}\mathbb{E}[u_2^2[k-1]])}_{:=\widetilde{P_2}}$$

Like the proof of Lemma 4.13, we denote the second and the third terms as $\widetilde{P_1}$ and $\widetilde{P_2}$ respectively.

• Power-Limited Interval: Denote

$$W_2 := a^{k-k_1-2}w[k_1+1] + \cdots + aw[k-2]$$
$$U_{12} := a^{k-k_1-2}u_1[k_1+1] + \cdots + u_1[k-1]$$
$$U_{22} := a^{k-k_1-2}u_2[k_1+1] + \cdots + u_2[k-1]$$

Here, U_{12} and U_{22} correspond to the first and second controller's input in the power-limited intervals described in Figure 4.23. We will first handle these power-limited inputs. Notice that

$$x[k] = a^{k-k_1-1}x[k_1+1] + a^{k-k_1-2}u_1[k_1+1] + \cdots + u_1[k-1] + a^{k-k_1-2}u_2[k_1+1] + \cdots + u_2[k-1]$$
$$+ a^{k-k_1-2}w[k_1+1] + \cdots + w[k-1].$$

Since $x[k_1+1]$ and W_2 are independent by causality, using Lemma 4.1 we can lower bound $\mathbb{E}[x^2[k]]$

as

$$
\begin{aligned}
&\mathbb{E}[x^2[k]] \\
&= \mathbb{E}[(a^{k-k_1-1}x[k_1+1] + U_{12} + U_{22} + W_2)^2] + 1 \\
&\geq (\sqrt{\mathbb{E}[(a^{k-k_1-1}x[k_1+1] + W_2)^2]} - \sqrt{\mathbb{E}[U_{12}^2]} - \sqrt{\mathbb{E}[U_{22}^2]})_+^2 + 1 \\
&= (\sqrt{\mathbb{E}[(a^{k-k_1-1}x[k_1+1])^2] + \mathbb{E}[W_2^2]} - \sqrt{\mathbb{E}[U_{12}^2]} - \sqrt{\mathbb{E}[U_{22}^2]})_+^2 + 1 \\
&\geq (\sqrt{a^{2(k-k_1-1)}\mathbb{E}[x[k_1+1]^2]} - \sqrt{\mathbb{E}[U_1^2]} - \sqrt{\mathbb{E}[U_2^2]})_+^2 + 1. \qquad (4.34)
\end{aligned}
$$

Here, $\mathbb{E}[x[k_1+1]^2]$ is lower bounded as

$$
\begin{aligned}
&\mathbb{E}[x[k_1+1]^2] = \mathbb{E}[(ax[k_1] + u_1[k_1] + u_2[k_1] + w[k_1])^2] \\
&= \mathbb{E}[(ax[k_1] + u_1[k_1] + u_2[k_1])^2] + \mathbb{E}[w[k_1]^2] \\
&\geq \mathbb{E}[(ax[k_1] + u_1[k_1] + u_2[k_1])^2]. \qquad (4.35)
\end{aligned}
$$

In the last term, the effect of the power-limited inputs is separated out.

• Information-Limited Interval: Using Lemma 4.9, we will bound the remaining uncertainty of the state after the information-limited interval. Since we will give all the disturbances except $w[0]$ as side-information, we denote the relevant observations as $y_1'[n]$ and $y_2'[n]$. Formally, denote

$$
\begin{aligned}
&W_1 := a^{k_1-1}w[0] + \cdots + w[k_1-1] \\
&U_{11} := a^{k_1-1}u_1[0] + \cdots + u_1[k_1-1] \\
&U_{21} := a^{k_1-1}u_2[0] + \cdots + u_2[k_1-1] \\
&\overline{W} := (w[1], \cdots, w[k-1]) \\
&y_1'[n] := a^{n-1}w[0] + v_1[n] \\
&y_2'[n] := a^{n-1}w[0] + v_2[n] \\
&W_1' := W_1 - \mathbb{E}[W_1 | y_1'[1:k_1-1], y_2'[1:k_1-1], \overline{W}] \\
&W_1'' := \mathbb{E}[W_1 | y_1'[1:k_1-1], y_2'[1:k_1-1], \overline{W}] \\
&u_1'[k_1] := u_1[k_1] - \mathbb{E}[u_1[k_1] | y_1'[1:k_1-1], y_2'[1:k_1-1], \overline{W}] \\
&u_1''[k_1] := \mathbb{E}[u_1[k_1] | y_1'[1:k_1-1], y_2'[1:k_1-1], \overline{W}] \\
&u_2'[k_1] := u_2[k_1] - \mathbb{E}[u_2[k_1] | y_1'[1:k_1-1], y_2'[1:k_1-1], \overline{W}] \\
&u_2''[k_1] := \mathbb{E}[u_2[k_1] | y_1'[1:k_1-1], y_2'[1:k_1-1], \overline{W}].
\end{aligned}
$$

Here, we have

$$
W_1' = a^{k_1-1}w[0] - \mathbb{E}[a^{k_1-1}w[0] | y_1'[1:k_1-1], y_2'[1:k_1-1]]
$$

Since $w[0], y_1'[1 : k_1-1], y_2'[1 : k_1-1], \overline{W}$ are jointly Gaussian, W_1' is independent from $y_1'[1 :$

$k_1-1], y'_2[1:k_1-1], \overline{W}$. By Lemma 4.9 we have

$$\mathbb{E}[W_1'^2] = \frac{a^{2(k_1-1)}\sigma_{v1}^2}{(1+\frac{\sigma_{v1}^2}{\sigma_{v2}^2})(\frac{a^{2(k_1-2)}(1-a^{-2(k_1-1)})}{1-a^{-2}}) + \sigma_{v1}^2} \tag{4.36}$$

This lower bounds the state disturbance due to $w[0]$ when it is observed by $y'_1[1:k_1-1]$ and $y'_2[1:k_1-1]$. Note that $y_1[1:k_1-1], y_2[1:k_1-1], \overline{W}$ is a function of $y'_1[1:k_1-1], y'_2[1:k_1-1], \overline{W}$. Therefore, U_{11} and U_{21} are also functions of $y'_1[1:k_1-1], y'_2[1:k_1-1], \overline{W}$ and (4.35) can be lower bounded as

$$\begin{aligned}
&\mathbb{E}[(ax[k_1]+u_1[k_1]+u_2[k_1])^2]\\
&=\mathbb{E}[(a(W_1+U_{11}+U_{12})+u_1[k_1]+u_2[k_1])^2]\\
&=\mathbb{E}[(aW'_1+u'_1[k_1]+u'_2[k_1])^2]+\mathbb{E}[(aW''_1+aU_{11}+aU_{12}+u''_1[k_1]+u''_2[k_1])^2]\\
&\geq \mathbb{E}[(aW'_1+u'_1[k_1]+u'_2[k_1])^2]
\end{aligned} \tag{4.37}$$

In the last term, the effect of the information-limited inputs is separated out.

• Radner's Interval: Now we will reduce the last term of (4.37) to Radner's problem. $u'_1[k_1]$ and $u'_2[k_1]$ are functions of $y_1[1:k_1], y'_1[1:k_1-1], y'_2[1:k_1-1], \overline{W}$ and $y_2[1:k_1], y'_1[1:k_1-1], y'_2[1:k_1-1], \overline{W}$ respectively. Here, $y'_1[1:k_1-1], y'_2[1:k_1-1], \overline{W}$ are independent from W'_1 and $y_1[1:k_1-1], y_2[1:k_1-1], \overline{W}$ is a function of $y'_1[1:k_1-1], y'_2[1:k_1-1], \overline{W}$. Therefore, only $y_1[k_1]$ at the first controller and $y_2[k_1]$ at the second controller are relevant to W'_1. Therefore, by removing independent parts from W'_1 in $y_1[k_1]$, the sufficient statistic of $y_1[k_1]$ is

$$\begin{aligned}
&y_1[k_1]-(w[k_1-1]+aw[k_1-2]+\cdots+a^{k_1-2}w[1])-\mathbb{E}[a^{k_1-1}w[0]|y'_1[1:k_1-1],y'_2[1:k_1-1]]\\
&-(u_1[k_1-1]+au_1[k_1-2]+\cdots+a^{k_1-2}u_1[1])\\
&-(u_2[k_1-1]+au_2[k_1-2]+\cdots+a^{k_1-2}u_2[1])\\
&=a^{k_1-1}w[0]-\mathbb{E}[a^{k_1-1}w[0]|y'_1[1:k_1-1],y'_2[1:k_1-1]]+v_1[k_1]\\
&=W'_1+v_1[k_1]
\end{aligned}$$

Likewise, $y_2[k_1]$ can be reduced to

$$W'_1+v_2[k_1]$$

Therefore, by considering W'_1 as an initial state, $v_1[k_1]$ and $v_2[k_1]$ as observation noises of the first and second controller, we can map the problem into Radner's problem. Here, we have the following power constraints on $u_1[k_1]$ and $u_2[k_1]$.

$$\begin{aligned}
(1-2.5^{-1})\mathbb{E}[u_1^2[k_1]] &\leq \widetilde{P_1}\\
(1-2.5^{-1})\mathbb{E}[u_2^2[k_1]] &\leq \widetilde{P_2}
\end{aligned}$$

Since a linear strategy is optimal in Radner's problem, by (4.36) we can conclude

$$\mathbb{E}[(aW_1' + u_1'[k_1] + u_2'[k_1])^2] \geq \inf_{c_1,c_2 \in \mathbb{R}} (a - c_1 - c_2)^2\Sigma + c_1^2\sigma_{v1}^2 + c_2^2\sigma_{v2}^2 \tag{4.38}$$

$$\text{s.t. } (1 - 2.5^{-1})c_1^2(\Sigma + \sigma_{v1}^2) \leq \widetilde{P_1}$$

$$(1 - 2.5^{-1})c_2^2(\Sigma + \sigma_{v2}^2) \leq \widetilde{P_2}$$

• Final Lower bound: Applying Lemma 4.10 with paramters $a = a$ and $b = 2.5^{-1}$, we can upper bound the power of the power-limited inputs.

$$\begin{aligned}\mathbb{E}[U_1^2] &= \mathbb{E}[(a^{k-k_1-2}u_1[k_1+1] + \cdots + u_1[k-1])^2] \\ &\leq \frac{a^{2(k-k_1-2)}(1-(2.5a^{-2})^{k-k_1-1})}{1-2.5a^{-2}} \frac{\widetilde{P_1}}{(1-2.5^{-1})2.5^{-1}}\end{aligned} \tag{4.39}$$

and likewise

$$\mathbb{E}[U_1^2] \leq \frac{a^{2(k-k_1-2)}(1-(2.5a^{-2})^{k-k_1-1})}{1-2.5a^{-2}} \frac{\widetilde{P_2}}{(1-2.5^{-1})2.5^{-1}} \tag{4.40}$$

Finally, plugging (4.38), (4.39), (4.40) into (4.34) gives the first bound based on $D_{L,2}(\widetilde{P_1}, \widetilde{P_2}; k_1, k, \Sigma)$:

$$\begin{aligned}\mathbb{E}[x^2[n]] \geq \inf_{c_1,c_2} (&\sqrt{a^{2(k-k_1-1)}((a-c_1-c_2)^2\Sigma + c_1^2\sigma_{v1}^2 + c_2^2\sigma_{v2}^2)} \\ &- \sqrt{\frac{a^{2(k-k_1-2)}(1-(2.5a^{-2})^{k-k_1-1})}{1-2.5a^{-2}} \frac{\widetilde{P_1}}{(1-2.5^{-1})2.5^{-1}}} \\ &- \sqrt{\frac{a^{2(k-k_1-2)}(1-(2.5a^{-2})^{k-k_1-1})}{1-2.5a^{-2}} \frac{\widetilde{P_2}}{(1-2.5^{-1})2.5^{-1}}})_+^2 + 1 \\ \text{s.t. } &(1-2.5^{-1})c_1^2(\Sigma + \sigma_{v1}^2) \leq \widetilde{P_1} \\ &(1-2.5^{-1})c_2^2(\Sigma + \sigma_{v2}^2) \leq \widetilde{P_2}\end{aligned}$$

The second bound based on $D_{L,3}(\widetilde{P_1}, \widetilde{P_2}; k_1)$ derived as follows. Since $\mathbb{E}[x^2[n]] \geq \mathbb{E}[w^2[n-1]] = 1$, trivially $D_L(\widetilde{P_1}, \widetilde{P_2}) \geq 1$. Moreover, as justified above, we have

$$\mathbb{E}[x^2[k_1]] \geq \mathbb{E}[(a^{k_1-1}w[0] - \mathbb{E}[a^{k_1-1}w[0]|y_1'[1:k_1-1], y_2'[1:k_1-1])^2] = \mathbb{E}[W_1'^2].$$

Therefore, by setting $k = k_1$ we get the second bound based on $D_{L,3}(\widetilde{P_1}, \widetilde{P_2}; k_1)$.

The last bound based on $D_{L,4}(\widetilde{P_1}, \widetilde{P_2}; k)$ of the lemma can be derived as follows.

$$\begin{aligned}&\mathbb{E}[x^2[k]] \\ &\geq (\sqrt{\mathbb{E}[(a^{k-1}w[0] + \cdots + w[k-1])^2]} - \sqrt{\mathbb{E}[(a^{k-1}u_1[0] + \cdots + u_1[k-1])^2]} \\ &- \sqrt{\mathbb{E}[(a^{k-1}u_2[0] + \cdots + u_2[k-1])^2]})_+^2 \\ &\geq (\sqrt{a^{2(k-1)}} - \sqrt{\frac{a^{2(k-2)}}{1-2.5a^{-2}} \frac{\widetilde{P_1}}{1-2.5^{-1}}} - \sqrt{\frac{a^{2(k-2)}}{1-2.5a^{-2}} \frac{\widetilde{P_2}}{1-2.5^{-1}}})_+^2\end{aligned}$$

where the first inequality follows from Lemma 4.1 and the second inequality follows from Lemma 4.10.

□

In this lemma, the time-interval from 0 to $k_1 - 1$ corresponds to the information-limited interval in Figure 4.23. The time-interval from k_1 to $k_1 + 1$ corresponds to the Radner's interval in Figure 4.23. The time-interval from $k_1 + 1$ to k corresponds to the power-limited interval in Figure 4.23.

4.7 Constant Ratio Optimality

Now, we have an upper and lower bound on $D(P_1, P_2)$. In this section, we will evaluate the bounds and prove Theorem 4.1 which bounds the weighted average cost within a constant ratio. Even though the numerical evaluations are not elegant[20], these are enough to justify constant ratio optimality.

The upper bounds are written from the power-disturbance tradeoff perspective of Problem F and denoted by $(D_U(P_1, P_2), P_1, P_2)$. The lower bounds in Lemma 4.12 and 4.13 are given for the original weighted average-cost of Problem B, which can be written as $(D_{L,i}(\widetilde{P_1}, \widetilde{P_2}), \widetilde{P_1}, \widetilde{P_2})$ from the power-disturbance perspective. The following lemma tells us that if these two tradeoff regions are within a constant ratio of each other as regions in $\mathbb{R}^3$, i.e. $\exists c \geq 1$ such that $(D_U(cP_1, cP_2), cP_1, cP_2) \leq c \cdot (D_{L,i}(P_1, P_2), P_1, P_2)$, then the average cost can be characterized to within a constant ratio.

Lemma 4.14. *For two functions $D_L(\widetilde{P_1}, \widetilde{P_2})$ and $D_U(P_1, P_2)$, let there exist $c \geq 1$ such that for all $x_1, x_2 \geq 0$*

$$D_U(cx_1, cx_2) \leq c \cdot D_L(x_1, x_2).$$

Then, for all $q, r_1, r_2 \geq 0$, the following inequality holds.

$$\min_{P_1, P_2 \geq 0} qD_U(P_1, P_2) + r_1 P_1 + r_2 P_2 \leq c(\min_{\widetilde{P_1}, \widetilde{P_2} \geq 0} qD_L(\widetilde{P_1}, \widetilde{P_2}) + r_1 \widetilde{P_1} + r_2 \widetilde{P_2})$$

Proof. Let $P_1^\star$ and $P_2^\star$ achieve the minimum of the right term of the inequality, i.e.

$$\begin{aligned}
&\min_{\widetilde{P_1}, \widetilde{P_2} \geq 0} qD_L(\widetilde{P_1}, \widetilde{P_2}) + r_1 \widetilde{P_1} + r_2 \widetilde{P_2} \\
&= qD_L(P_1^\star, P_2^\star) + r_1 P_1^\star + r_2 P_2^\star.
\end{aligned}$$

[20]The bounds can probably be improved and tightened. However, the main concern of this chapter is not quantifying the exact cost, but qualitatively understanding the near-optimal strategies. The constant ratio optimality results are enough to justify our intuition behind the proposed strategies.

Then, we have

$$
\begin{aligned}
&c(\min_{\widetilde{P_1},\widetilde{P_2}\geq 0} qD_L(\widetilde{P_1},\widetilde{P_2}) + r_1\widetilde{P_1} + r_2\widetilde{P_2}) \\
&= c \cdot (qD_L(P_1^\star, P_2^\star) + r_1 P_1^\star + r_2 P_2^\star) \\
&\geq qD_U(cP_1^\star, cP_2^\star) + r_1(cP_1^\star) + r_2(cP_2^\star) \\
&\geq \min_{P_1,P_2} qD_U(P_1,P_2) + r_1P_1 + r_2P_2
\end{aligned}
$$

where the first inequality comes from the assumption of the lemma. Thus, the lemma is proved. □

We will show that the proposed strategies of Defintion 4.1 and 4.2 solve the weighted average cost problem of Problem B to within a constant ratio. Let's call the case when $\sigma_{v2}^2 \leq \max(1, a^2\sigma_{v1}^2)$ the *weakly-degraded-observation* case since the gap between the two controllers' observation noises is not too huge and the second controller can observe what the first controller observed only after one-time step. Likewise, we will call the case when $\sigma_{v2}^2 > \max(1, a^2\sigma_{v1}^2)$ the *strongly-degraded-observation* case since the gap between the observation noises is larger.

The weakly-degraded-observation case will be discussed in Section 4.7.1 and the strongly-degraded-observation case will be covered in Section 4.7.2.

4.7.1 Weakly-Degraded-Observation case with $|a| \geq 2.5$

Let's first consider the weakly-degraded case when $\sigma_{v2}^2 \leq \max(1, a^2\sigma_{v1}^2)$, which corresponds to the left half plane of Figure 4.9 of page 164. In this case, the infeasibility of 0-stage signaling discussed in Section 4.3.1 and 4.6.6 shows up and thus linear strategies are enough for constant-ratio optimality.

First, we evaluate the lower bound of Lemma 4.13 which involves Radner's problem.

Corollary 4.1. *Let $|a| \geq 2.5$ and $\sigma_{v2}^2 \leq \max(1, a^2\sigma_{v1}^2)$. Then, for all $q, r_1, r_2 \geq 0$, the minimum cost* (4.1) *of Problem B is lower bounded as follows:*

$$
\inf_{u_1,u_2} \limsup_{N\to\infty} \frac{1}{N} \sum_{0\leq n<N} q\mathbb{E}[x^2[n]] + r_1\mathbb{E}[u_1^2[n]] + r_2\mathbb{E}[u_2^2[n]] \geq \min_{\widetilde{P_1},\widetilde{P_2}\geq 0} qD_L(\widetilde{P_1},\widetilde{P_2}) + r_1\widetilde{P_1} + r_2\widetilde{P_2}
$$

where $D_L(\widetilde{P_1},\widetilde{P_2})$ satisfies the following conditions.
(a) If $\widetilde{P_1} \leq \frac{1}{400}a^2\max(1, a^2\sigma_{v1}^2)$ and $\widetilde{P_2} \leq \frac{1}{400}a^2\max(1, a^2\sigma_{v2}^2)$ then $D_L(\widetilde{P_1},\widetilde{P_2}) = \infty$.
(b) If $\widetilde{P_1} \leq \frac{1}{400}a^2\max(1, a^2\sigma_{v1}^2)$, for all $\widetilde{P_2}$, $D_L(\widetilde{P_1},\widetilde{P_2}) \geq 0.176a^2\sigma_{v2}^2 + 1$.
(c) For all $\widetilde{P_1}$ and $\widetilde{P_2}$, $D_L(\widetilde{P_1},\widetilde{P_2}) \geq 0.295 \cdot \max(1, a^2\sigma_{v1}^2)$.

Proof. See Appendix 9.1. □

(a) and (b) tell what happens if the first controller has little power (i.e. it must follow something close to a zero-input strategy). (a) shows if the second controller does not have enough power, the system becomes unstable. (b) shows that even if the second controller has enough power, the state variance is lower bounded by the second controller's observation noise. (c) shows that even if the first controller has enough power to apply a zero-forcing strategy, the state variance is lower bounded by the first controller's observation noise.

The following lemma analyzes the achievable disturbance by the simple linear strategy of Definition 4.1.

Lemma 4.15. *Consider a single-controller scalar system*

$$x[n+1] = ax[n] + u[n] + w[n]$$
$$y[n] = x[n] + v[n]$$

where $w[n]$ *is i.i.d.* $\mathcal{N}(0,1)$ *and* $v[n]$ *is i.i.d.* $\mathcal{N}(0,\sigma_v^2)$. *For a given control strategy, let* $D(P) := \limsup_{n\to\infty} \frac{1}{N}\sum_{0\le n<N} \mathbb{E}[x^2[n]]$ *and* $P := \limsup_{n\to\infty} \frac{1}{N}\sum_{0\le n<N} \mathbb{E}[u^2[n]]$. *Then,*

$$(D(P), P) \le (a^2\sigma_v^2 + 1, a^4\sigma_v^2 + a^2\sigma_v^2 + a^2)$$

is achievable by a linear bang-bang controller, $u[n] = -ay[n]$. *Therefore, in Problem F the following power-disturbance tradeoffs are achievable.*

$$(D(P_1,P_2), P_1, P_2) \le (a^2\sigma_{v1}^2 + 1, a^4\sigma_{v1}^2 + a^2\sigma_{v1}^2 + a^2, 0),$$
$$(D(P_1,P_2), P_1, P_2) \le (a^2\sigma_{v2}^2 + 1, 0, a^4\sigma_{v2}^2 + a^2\sigma_{v2}^2 + a^2).$$

Proof. Put $u[n] = -ay[n]$ into the system equation. Then, we have

$$x[n+1] = ax[n] - ax[n] - av[n] + w[n]$$
$$= -av[n] + w[n]$$

Thus, we conclude for $n \ge 1$

$$\mathbb{E}[x^2[n]] = a^2\sigma_v^2 + 1$$

and

$$\mathbb{E}[u^2[n]] = a^2\mathbb{E}[(x[n]+v[n])^2]$$
$$= a^2(a^2\sigma_v^2 + 1 + \sigma_v^2)$$

□

Using Lemma 4.14, Corollary 4.1 and Lemma 4.15, we can compare the upper and lower bounds to prove linear strategies suffice to achieve constant-ratio optimality in this region of problem parameters.

Proposition 4.7. *There exists $c \geq 1200$ such that for all $a, q, r_1, r_2, \sigma_0, \sigma_{v1}, \sigma_{v2}$ satisfying $|a| \geq 2.5$ and $\sigma_{v2}^2 \leq \max(1, a^2\sigma_{v1}^2)$, the following inequality holds:*

$$\frac{\inf_{u_1,u_2 \in L_{lin,bb}} \limsup_{N\to\infty} \frac{1}{N}\sum_{0\leq n<N} \mathbb{E}[qx^2[n] + r_1u_1^2[n] + r_2u_2^2[n]]}{\inf_{u_1,u_2} \limsup_{N\to\infty} \frac{1}{N}\sum_{0\leq n<N} \mathbb{E}[qx^2[n] + r_1u_1^2[n] + r_2u_2^2[n]]} \leq c.$$

Proof. See Appendix 9.2. □

4.7.2 Strongly-Degraded-Observation case with $|a| \geq 2.5$

Let's consider the strongly-degraded-observation case when $\sigma_{v2}^2 > \max(1, a^2\sigma_{v1}^2)$, which corresponds to the right half-plane of Figure 4.16. Since $|a| \geq 2.5$, we can find $s \in \mathbb{N}$ such that $a^{2(s-1)}\max(1, a^2\sigma_{v1}^2) \leq \sigma_{v2}^2 \leq a^{2s}\max(1, a^2\sigma_{v1}^2)$. We will show that the s-stage signaling strategy is required for constant-ratio optimality.

Since we need a matching lower bound to s-stage signaling strategies, we evaluate Lemma 4.13 which has a generalized Witsenhausen's counterexample in it.

Corollary 4.2. *Let $|a| \geq 2.5$ and for some $s \in \mathbb{N}$, suppose*

$$a^{2(s-1)}\max(1, a^2\sigma_{v1}^2) \leq \sigma_{v2}^2 \leq a^{2s}\max(1, a^2\sigma_{v1}^2).$$

Then, for all $q, r_1, r_2 \geq 0$, the minimum cost (4.1) of Problem B is lower bounded as follows:

$$\inf_{u_1,u_2} \limsup_{N\to\infty} \frac{1}{N} \sum_{0\leq n<N} q\mathbb{E}[x^2[n]] + r_1\mathbb{E}[u_1^2[n]] + r_2\mathbb{E}[u_2^2[n]] \geq \min_{\widetilde{P_1},\widetilde{P_2}\geq 0} qD_L(\widetilde{P_1},\widetilde{P_2}) + r_1\widetilde{P_1} + r_2\widetilde{P_2}.$$

where $D_L(\widetilde{P_1},\widetilde{P_2})$ satisfies the following conditions.
(a) When $\widetilde{P_1} \leq \frac{\sigma_{v2}^2}{70a^{2(s-1)}}$, then $D_L(\widetilde{P_1},\widetilde{P_2}) \geq 0.008a^2\sigma_{v2}^2 + 1$.
(b) When $\widetilde{P_1} \leq \frac{\sigma_{v2}^2}{70a^{2(s-1)}}$ and $\widetilde{P_2} \leq \frac{a^4\sigma_{v2}^2}{28000}$, then $D_L(\widetilde{P_1},\widetilde{P_2}) = \infty$.
(c) When $\frac{\sigma_{v2}^2}{70a^{2(s-1)}} \leq \widetilde{P_1} \leq \frac{1}{20000}\max(a^2, a^4\sigma_{v1}^2)$,
then $D_L(\widetilde{P_1},\widetilde{P_2}) \geq 0.2541a^{2s}\widetilde{P_1}\exp(-\frac{50a^{2(s-1)}\widetilde{P_1}}{\sigma_{v2}^2}) + 0.066a^{2s}\max(1, a^2\sigma_{v1}^2) + 1$.
(d) When $\frac{\sigma_{v2}^2}{70a^{2(s-1)}} \leq \widetilde{P_1} \leq \frac{1}{20000}\max(a^2, a^4\sigma_{v1}^2)$ and $\widetilde{P_2} \leq 0.0457a^{2(s+1)}\widetilde{P_1}\exp(-\frac{50a^{2(s-1)}\widetilde{P_1}}{\sigma_{v2}^2}) + 0.0113a^{2(s+1)}\max(1, a^2\sigma_{v1}^2)$, then $D_L(\widetilde{P_1},\widetilde{P_2}) = \infty$.
(e) For all $\widetilde{P_1}$ and $\widetilde{P_2}$, $D_L(\widetilde{P_1},\widetilde{P_2}) \geq 0.295 \cdot \max(1, a^2\sigma_{v1}^2)$.

Proof. See Appendix 9.3. □

(a) and (b) tell what happens if the first controller has little power and thus is forced to be close to a zero-input strategy. Even if the second controller has enough power, the state variance is lower bounded by the second controller's observation noise. If the second controller does not have enough power to stabilize the system, the state diverges to infinity. (e) shows the opposite case when the first controller has enough power to apply a zero-forcing strategy. However, even in this case, the state variance is lower bounded by the first controller's observation noise. (c) and (d) cover the cases between these two extreme cases. (c) gives the lower bound that matches to the s-stage signaling strategy when the second controller has enough power. (d) shows that since the first controller does not stabilize the system with its signaling alone, the second controller's input power has to be large enough to stabilize the system.

Now, we evaluate the performance of the s-stage signaling analyzed in Lemma 4.7 of page 174.

Corollary 4.3. *Consider Problem F of page 163, and let* $|a| \geq 2.5$ *and suppose* $a^{2(s-1)}\max(1, a^2\sigma_{v1}^2) \leq \sigma_{v2}^2 \leq a^{2s}\max(1, a^2\sigma_{v1}^2)$ *for some* $s \in \mathbb{N}$*. Then, there exists an upper bound* $D_U(P_1, P_2)$ *on* $D(P_1, P_2)$ *i.e.* $D(P_1, P_2) \leq D_U(P_1, P_2)$ *for all* $P_1, P_2 \geq 0$ *satisfying the following:*

$$\begin{aligned}(D_U(P_1, P_2), P_1, P_2) \leq &(832a^{2s}P\exp(-\frac{50a^{2(s-1)}P}{\sigma_{v2}^2}) + 63a^{2s}\max(1, a^2\sigma_{v1}^2),\\ &80000P, 6656a^{2(s+1)}P\exp(-\frac{50a^{2(s-1)}P}{\sigma_{v2}^2}) + 564a^{2(s+1)}\max(1, a^2\sigma_{v1}^2))\end{aligned}$$

for $\frac{\sigma_{v2}^2}{70a^{2(s-1)}} \leq P \leq \frac{1}{20000}\max(a^2, a^4\sigma_{v1}^2)$.

Proof. See Appendix 9.4 □

Here we can notice that the performance is matching that of Corollary 4.2 (c), (d) in that the bounds on the state disturbance take the same form of a function on P_1 and system parameters.

Now, we can compare these two bounds to prove constant-ratio optimality.

Proposition 4.8. *There exists* $c \leq 1.5 \times 10^5$ *such that for all* $a, q, r_1, r_2, \sigma_0, \sigma_{v1}, \sigma_{v2}$ *satisfying* $|a| \geq 2.5$ *and*

$$a^{2(s-1)}\max(1, a^2\sigma_{v1}^2) \leq \sigma_{v2}^2 \leq a^{2s}\max(1, a^2\sigma_{v1}^2).$$

for some $s \in \mathbb{N}$*, the following inequality holds:*

$$\frac{\inf_{u_1,u_2 \in L_{lin,bb} \cup L_{sig,s}} \limsup_{N\to\infty} \frac{1}{N}\sum_{0\leq n<N} \mathbb{E}[qx^2[n] + r_1u_1^2[n] + r_2u_2^2[n]]}{\inf_{u_1,u_2} \limsup_{N\to\infty} \frac{1}{N}\sum_{0\leq n<N} \mathbb{E}[qx^2[n] + r_1u_1^2[n] + r_2u_2^2[n]]} \leq c.$$

Proof. See Appendix 9.5 □

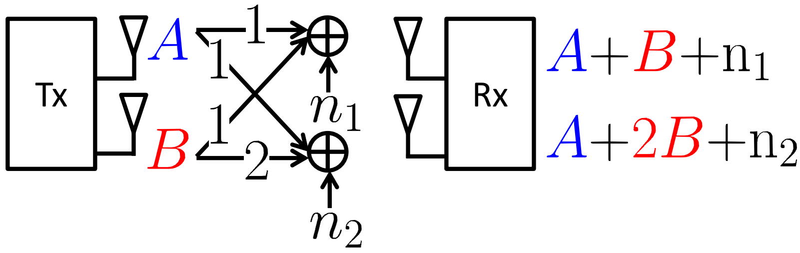
Now, Theorem 4.1 immediately follows from Propositions 4.7 and 4.8.

Proof of Theorem 4.1. Propostion 4.7 covers the case when $\sigma_{v2}^2 \leq \max(1, a^2\sigma_{v1}^2)$. Propostion 4.8 covers the case when $\sigma_{v2}^2 > \max(1, a^2\sigma_{v1}^2)$, since in this case there exists $s \in \mathbb{N}$ such that

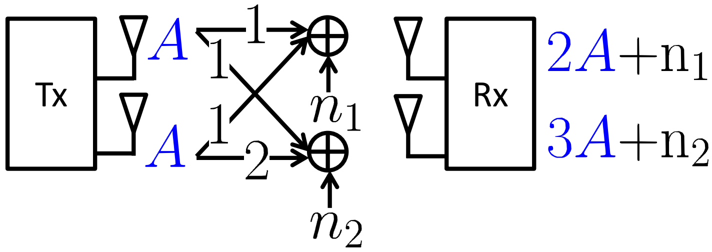
$$a^{2(s-1)}\max(1, a^2\sigma_{v1}^2) \leq \sigma_{v2}^2 \leq a^{2s}\max(1, a^2\sigma_{v1}^2).$$

□

4.8 Connection to Wireless Communication Theory



(a)



(b)

Figure 4.24: MIMO Wireless Communication Problem: (a) By transmitting different signals across the antennas, we can achieve 'd.o.f. gain'. Generally, this scheme performs well in high-SNR. (b) By transmitting the same signal across the antennas, we can achieve 'power gain'. Generally, this scheme performs well in low-SNR.

Throughout the discussion, we have observed a lot of similarity between wireless communication and the decentralized LQG control problems considered in this chapter. In this section, we will explore this point in more detail. At first glance, decentralized LQG control and wireless communication seem pretty distinct from each other. But the main result in this chapter is actually a manifestation of a deeper connection.

The essence of wireless communication problems [99] can be summarized as follows: First, unlike wired communication, wireless communication systems share a common channel and as a result the signals from different transmitting antennas can interact with each other. Second, wireless communication systems involve uncertainties or randomness that come from channel fading or thermal noise in circuits. Third, to extend battery life and minimize interference to other transceivers, each transmitting antenna has a power constraint.

The way that wireless communication theory models capture this nature is very similar to stochastic control theory. First, the interaction between signals is modeled by linear operations. Second, the uncertainty in the system is modeled by Gaussian random variables. Third, the power of the transmitting antennas is measured by a quadratic cost. If we remember that wireless communication systems are by nature distributed, wireless communication problems are essentially a special case of decentralized LQG control problems, except that wireless communication problems have the special objective of communication.

Like decentralized LQG problems, wireless network communication problems are still open [29, 6] or nonconvex [115]. However, wireless communication theorists found that it is helpful to divide cases according to the SNR(Signal-to-Noise Ratio). For a given communication scheme, the capacity of a channel is usually given as $\log(1 + c_1 SNR + \cdots + c_k SNR^k)$. Therefore, when SNR is large (high-SNR case), the capacity is approximately $k \log SNR$ (where k turns out to be the 'd.o.f. gain' of the scheme). When SNR is small (low-SNR case), the capacity is approximately $c_1 SNR$ (where c_1 turns out to be the 'power gain' of the scheme). Therefore, depending on the SNR the capacity of communication schemes are very different. Consequently, wireless communication theory usually divides into two cases: (1) high-SNR (2) low-SNR.

Let's consider the 2×2 MIMO communication problem shown in Figure 4.24. We can think of two basic ways of exploiting these antennas. The first way is transmitting different signals across different antennas. As we can see in Fig. 4.24a, in this case the receiver will have two variables and two (noisy) equations, and we can expect 'MIMO gain' by solving for multiple variables. In wireless communication theory, this gain is called the 'd.o.f.(degree-of-freedom) gain' and the scheme of Figure 4.24a succeeds in increasing k in the capacity formula.

As we mentioned in Section 4.5.1, this concept can be extended to *generalized d.o.f.* by allowing the transmitting powers of different antennas to scale differently [29]. When the transmitting powers of different antennas scale differently, we can further divide a single receiving antenna according to "signal levels". For the small signal level, all tranmitting antennas can affect it, but for the large signal level, only the few transmitting antennas with large power can affect it. In [6], binary deterministic models were proposed to capture this phenomenon by conceptualizing different bit-levels like different antennas, which we used in Section 4.3.

The second way of using the antennas is transmitting the same signal across different antennas as shown in Fig. 4.24b. In this scheme, the receiver will have only one variable and we cannot expect the d.o.f. gain of solving for multiple variables. However, there is a gain to be had from aligning the signals. Let's assume all random variables, A, B, n_1, n_2, are Gaussian random variables with zero mean and unit variance, and compute the signal-to-noise ratio at the receiving antennas. The SNR of the first receive antenna in Fig. 4.24a is $\frac{\mathbb{E}[(A+B)^2]}{\mathbb{E}[n_1^2]} = 2$. On the other hand, the SNR of the first antenna in Fig. 4.24b is $\frac{\mathbb{E}[(2A)^2]}{\mathbb{E}[n_1^2]} = 4$. Therefore, by transmitting the same signal over different antennas we can increase the SNR of the received signals. This gain is known as 'power gain' in wireless communication theory and the proposed scheme is good for increasing c_1 in the capacity formula. To exploit the power gain, the receiver has to introduce maximum-ratio combining [99].

How is this relevant for scalar decentralized LQG control problems? To control a plant we first have to gain information about its state. The quantitative behavior of this information

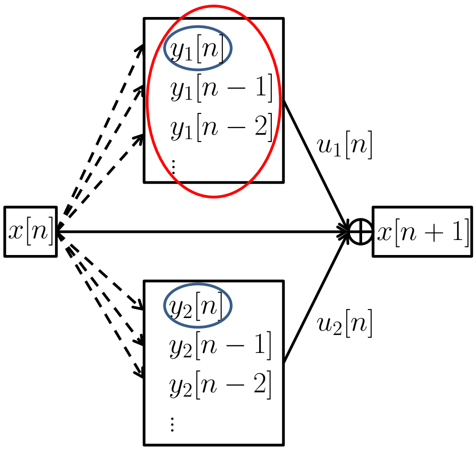


Figure 4.25: As indicated by the blue circles, in the fast dynamics case the proposed scheme exploits "information flow" from both controllers but ignores all past observations. By contrast, as indicated by the red circle, in the slow dynamics case the proposed scheme exploits the information of only one controller but takes into account all past observations.

flow (from a plant to controllers and finally back to the plant as we discussed in Chapter 3) is very similar to that in wireless communication systems. More precisely, according to the **eigenvalue** of the system, the information flow in the system shows a very different behavior. The system is deemed to be fast-dynamics when the eigenvalue is large ($|a| \geq 2.5$) and slow-dynamics when the eigenvalue is small ($|a| < 2.5$).

The main reason for this division is the relationship between the eigenvalue of the system and the SNR of the information flow for control. The discussion of Section 4.3 reveals that the SNR of implicit communication between two controllers will be bounded[21] by the eigenvalue squared ($|a|^2$). Therefore, when the eigenvalue is large, the SNR for implicit communication is also large. Therefore, from wireless communication theory we can expect that the (generalized) d.o.f. gain of the implicit communication is crucial. Likewise, when the eigenvalue is small, the SNR for implicit communication is also small and the power gain of the implicit communication is crucial. This is the slow dynamics case. To harness the power gain, we have to use Kalman filtering which corresponds to maximum-ratio combining in wireless communication [99].

In short, even if the system is the simplest scalar system, we can think of two ways of sending information. One way is across different bit-levels and the other way is across different time-slots. Moreover, these multiple bit-levels and multiple time-slots can be deemed as MIMO antennas in wireless communication theory. In fast-dynamics, the MIMO antenna gain of multiple

[21]Since the second controller can cancel all bits above its noise level at the next time step, the new information in the state cannot be amplified more than $|a|^2$ above the second controller's noise level. Thus, the SNR measured at the second controller is always bounded by $|a|^2$.

bit-levels dominates that of multiple time-slots. The proposed signaling strategies exploit the d.o.f. gain of the MIMO antennas over multiple bit-levels.

On the other hand, in slow-dynamics, the MIMO antenna gain of multiple time-slots is much more crucial. In Chapter 5, Kalman filtering will be used to exploit the power gain of the MIMO antennas over multiple time-slots.

Figure 4.25 visualizes the discussion so far. In fast-dynamics, the state is quickly changing and the SNR of implicit communication is high. Thus, the information from previous time steps is much less important than that of the current time step. However, to fully exploit the MIMO antenna gain of different bit-levels, the observations from both controllers have to be used.[22] On the other hand, in slow-dynamics, the state changes slowly and the SNR of implicit communication is low. Therefore, there is no huge incentive for implicit communication between controllers, and a strategy which fully exploits the observations of either one controller is enough to achieve constant-ratio optimality. However, the power gain from the past observations cannot be ignored and so Kalman filtering has to be used.

It is worth mentioning that this fundamental difference between fast and slow dynamics was conjectured as early as the 1970s [88] but remained vague: *"The development of systematic procedures for appropriately modeling large scale systems with slow and fast dynamics has not received the attention it deserves. ⋯ one should look for time scale separation (fast and slow dynamics)."* This dissertation is the first that has used this quantitatively.

The division of fast and slow dynamics based on 2.5 is somewhat surprising if we recall that Witsenhausen's counterexample seems to correspond to the $a = 1$ case in the infinite horizon problem. In [37] it was shown that we need a nonlinear strategy to achieve constant-ratio optimality in Witsenhasuen's counterexample, while in Chapter 5 it is shown that in the slow dynamics case (including $a = 1$) linear strategies are enough for a constant ratio optimality. The main reason for this is that the infinite horizon problem is a sequential problem. Since the problem is sequential, we can think of the infinite-horizon problem as an interlocking series of Witsenhausen's counterexamples. When $a = 1$, the interference from the previous Witsenhausen's problem dominates and we do not have to solve the current Witsenhausen's problem optimally.

4.9 Discussion and Further Research

In the beginning of the chapter, we summarized the two main contributions of the classical centralized LQG result. The first was linear controller optimality which narrows the search for the optimal strategy from the infinite-dimensional strategy space to the finite-dimensional linear strategy space. In Theorem 4.1, we gave the corresponding result for scalar decentralized LQG problems in

[22] Even though the strategy in Definition 4.2 relies on the past controller inputs (therefore, the past observations), the role of the past inputs in the strategy is just to cancel their influence on the current time step, not providing information about the state.

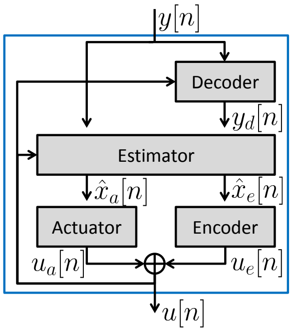


Figure 4.26: Communication-Estimation-Control Separation Controllers

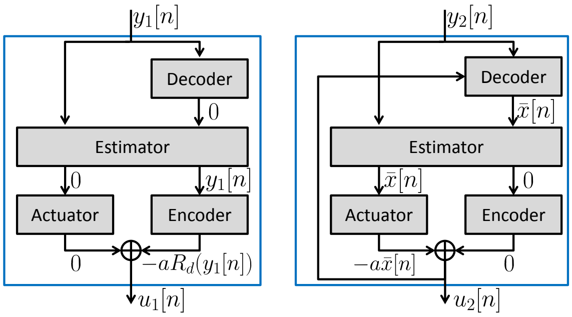


Figure 4.27: Communication-Estimation-Control separation controller interpretation of $L_{sig,s}$ (Here, $\bar{x}[n] = Q_{a^s d}(y_2[n] - R_{a^s d}(\sum_{1 \le i \le s} a^{i-1} u_2[n-i])) + R_{a^s d}(\sum_{1 \le i \le s} a^{i-1} u_2[n-i]))$

an approximation sense by proposing a finite-dimensional strategy space.

The second contribution was more philosophical. Centralized LQG gave us the separation principle for estimation and control. Therefore, a natural question is whether we can interpret the result of Theorem 4.1 in terms of estimation-control separation, or if there is a conceptual missing block. The authors believe there is a missing fundamental design block, the "communication" block.

Figure 4.26 shows the proposed communication-estimation-control separation controller, which we believe, is **approximately** optimal. First, the controller observes $y[n]$. Unlike the centralized case, $y[n]$ may contain transmitted "signals" from the other controllers. The decoder block extracts such information and generates a new observation $y_d[n]$. Based on both $y[n]$ and $y_d[n]$, the estimator block tries to estimate the states. After the estimation, the controller can either control the states[23] by itself, or relay information to the other controllers and let them control. $\hat{x}_a[n]$ is the states that the controller wants to control by itself. Based on $\hat{x}_a[n]$, the actuator generates the control action $u_a[n]$. $\hat{x}_e[n]$ is the state that the controller wants to encode for the other controllers. Based on $\hat{x}_e[n]$, the encoder generates the encoded signal $u_e[n]$. Finally, the control output is the superposition of $u_a[n]$ and $u_e[n]$.

Figure 4.27 interprets the strategy $L_{sig,s}$ based on the proposed controller structure. The strategy exploits the fact that the controller 1 has a better observation than the controller 2. Since the controller 1's control signal is expensive, it "relays" its observation through the encoder rather than trying to control the state by itself. Then, the controller 2 extracts the relayed information in the decoder block, and takes action based on it. We can notice that only encoders and decoders are nonlinear, while estimators and actuators are linear. Therefore, this structure fits the intuition that the essential nonlinearity comes from communication.

An extensive relationship between the implicit information flow for control and wireless information flow was discussed in Section 4.8. Some unique features of information flows for control were also noticed. We also found the counterpart of the classical notion of information theoretic cutset bounds [21] in the dynamic-programming context. The geometric slicing idea discussed in Figure 4.14 can be thought of as a cutset bound in a sense that it finds the informational bottleneck of the system. However, unlike traditional information-theoretic cutsets, the geometric slicing idea divides the nodes by a weighted cut rather than a simple partitioning.

Even if this chapter focused on the simplest toy scalar LQG problem with two controllers, the essential difficulty of decentralized problems — nonconvex optimization over infinite-dimensional space — was still there and we could finesse this difficulty by taking an approximation approach. We believe the approaches and techniques developed in this chapter will also be useful in more general problems with vector states and multiple controllers. Moreover, in the process of such generalization, we will find more close relationships and parallels between wireless information flows and control information flows. For example, the notion of the computation over communication channels [74]

[23] Even a single scalar state can be viewed as a collection of bit-positions.

or interference alignment [63, 14] has yet to be properly understood in control contexts. Above all, by solving the problems only approximately, we “may” be able to make a breakthrough in this long-standing open problem, the decentralized LQG problem.

Chapter 5

Decentralized scalar LQG problem: Slow Dynamics

5.1 Introduction

In Chapter 4, we consider the simplest decentralized LQG (linear quadratic Gaussian) problem, the scalar infinite-horizon LQG problem with two controllers. In the last chapter, we focused on the fast dynamics case when the eigenvalue of the system is large. The most interesting fact in this case is that a nonlinear control strategy can infinitely outperform any linear strategy especially when the two controllers are asymmetric. When the first controller has a better observation with high control cost and the second controller has a worse observation with small control cost, there is a huge incentive for the first controller to communicate its observation to the second controller. Moreover, this communication is implicitly through the plant and for such implicit communication, nonlinear strategies are more efficient than linear strategies. The Signal-to-Noise Ratio (SNR) for this implicit communication is upper bounded by the eigenvalue of the system. Thus, as the eigenvalue of the system goes to infinity, the performance gap between nonlinear and linear strategy can unboundedly diverge.

In this chapter, we focus on slow dynamics where the eigenvalue of the system is bounded by a constant. The SNR for implicit communication in this case is bounded and the performance gap between the best nonlinear and linear strategies is bounded by a constant. In the scalar system considered in this chapter, the system is observable and controllable by both controllers. It turns out that control by a single controller is good enough to achieve a constant-ratio of the optimal cost.

The rest of the chapter consists as follows: In Section 5.2, we will state the problem and main results. In Section 5.3, we will revisit classic centralized control results and intuitively understand them. In Section 5.4, we will derive a fundamental lower bound on the control performance, and prove that the centralized control performance and the derived lower bound are within a constant ratio.

5.2 Problem Statement and Main Result

Throughout this chapter, we will consider the same problem considered in Chapter 4, the scalar infinite-horizon decentralized LQG problems with two controllers. However, while the focus of Chapter 4 was the fast-dynamics case (when $|a| \geq 2.5$), the focus of this chapter is the slow-dynamics case (when $|a| < 2.5$).

Problem J (scalar infinite-horizon decentralized LQG problems with two controllers)**.** *Consider the system dynamics given as*

$$\begin{aligned} x[n+1] &= ax[n] + u_1[n] + u_2[n] + w[n] \\ y_1[n] &= x[n] + v_1[n] \\ y_2[n] &= x[n] + v_2[n] \end{aligned}$$

where $x[0] \sim \mathcal{N}(0, \sigma_0^2)$*,* $w[n] \sim \mathcal{N}(0,1)$*,* $v_1[n] \sim \mathcal{N}(0, \sigma_{v1}^2)$*,* $v_2[n] \sim \mathcal{N}(0, \sigma_{v2}^2)$ *are independent Gaussian random variables. The control inputs,* $u_1[n]$ *and* $u_2[n]$*, must be causal functions of* $y_1[n]$ *and* $y_2[n]$ *respectively, i.e.* $u_1[n] = f_{1,n}(y_1[0], \cdots, y_1[n])$ *and* $u_2[n] = f_{2,n}(y_2[0], \cdots, y_2[n])$*.*

For $q, r_1, r_2 \geq 0$*, the control objective is to minimize a long-term average quadratic cost:*

$$\limsup_{N \to \infty} \frac{1}{N} \sum_{0 \leq n < N} q\mathbb{E}[x^2[n]] + r_1 \mathbb{E}[u_1^2[n]] + r_2 \mathbb{E}[u_2^2[n]]. \tag{5.1}$$

As discussed in Chapter 4, even though we normalized the problem parameters (the variance of $w[n]$, the gains for $u_1[n], u_2[n], y_1[n], y_2[n]$), this problem includes all scalar two-controller decentralized LQG problems by a proper scaling.

In Chapter 4, we saw that in fast-dynamics cases, implicit communication between two controllers is crucial to achieve the optimal performance within a constant ratio. Moreover, essentially memoryless controllers, which only exploit the information at the current time step, were constant-ratio optimal.

Therefore, a natural question for slow-dynamics cases is that whether the same type of controllers are enough to achieve constant ratio optimality. In other words, is implicit communication crucial for performance? Can memoryless controllers achieve constant-ratio optimality? In this chapter, we will see that the answers for both questions are negative.

To understand why the answer for the first question is negative, let's revisit fast-dynamics cases. Even though the mathematical definition of implicit communication is still unclear, we can roughly measure the SNR (signal-to-noise ratio) of implicit communication. The blurry controller (the controller with higher observation noise) utilizes the transmitted signal from the other controller as soon as the transmitted signal's power exceeds its observation noise level. Therefore, the maximum SNR for implicit communication cannot exceed a^2, which is the ratio at which the system dynamics amplify signals in one time step. From this, we can conjecture that for slow dynamics cases ($|a| \leq 2.5$), the SNR is bounded and implicit communication may not be crucial for constant-ratio optimality.

However, justification is not that simple since the time-horizon is infinite. In other words, even though we could justify that the SNR at each time step is bounded, accumulation of such information may result in unbounded gain. Furthermore, a precise definition of implicit communication and the corresponding SNR requires further study.

For the second question, we will see that all observations from the past have to be utilized to achieve constant-ratio optimality. For this, Kalman filtering must be used.

In other words, we will prove that in the slow-dynamics case, single-controller optimal strategies — Kalman filtering linear strategies — are approximately optimal within a constant ratio. For this, let's first define the single-controller strategies which involve only one parameter k.

Definition 5.1 (Single Controller Optimal Strategy $L_{lin,kal}$). *$L_{lin,kal}$ is the set of all controllers which can be written in either one of two following forms for some $k \in \mathbb{R}$*
(i) $u_1[n] = -k\mathbb{E}[x[n]|y_1^n, u_1^{n-1}]$, $u_2[n] = 0$
(ii) $u_1[n] = 0$, $u_2[n] = -k\mathbb{E}[x[n]|y_2^n, u_2^{n-1}]$

Here, we can notice that since the system is linear and underlying random variables are Gaussian, the conditional expectations are linear in the observations [11].

Now, we can state the main theorem of this chapter, which states that when $|a| \leq 2.5$ the optimization only over $L_{lin,kal}$ is enough to achieve approximate optimality within a constant ratio among all possible strategies.

Theorem 5.1. *Consider the decentralized LQG problem shown in Problem J. Let L be the set of all measurable causal strategies. Then, there exists a constant $c \leq 2 \cdot 10^6$ such that for all $|a| \leq 2.5$, q, r_1, r_2, σ_0, σ_{v1} and σ_{v2},*

$$\frac{\inf_{u_1,u_2 \in L_{lin,kal}} \limsup_{N \to \infty} \frac{1}{N} \sum_{0 \leq n < N} \mathbb{E}[qx^2[n] + r_1 u_1^2[n] + r_2 u_2^2[n]]}{\inf_{u_1,u_2 \in L} \limsup_{N \to \infty} \frac{1}{N} \sum_{0 \leq n < N} \mathbb{E}[qx^2[n] + r_1 u_1^2[n] + r_2 u_2^2[n]]} \leq c$$

Proof. See Section 5.4.5 for the proof. □

The basic proof strategy is following. Rather than directly considering the average cost problem of Problem J, we consider the power-distortion tradeoff problem of Problem K. Then, since we have an explicit constraint on the controller power, we can divide the tradeoff curve into multiple regions based on the control power. For these finite number of regions, we derive different upper and lower bounds on the performance. By comparing them, we characterize the tradeoff curve within a constant ratio. Then, we finally convert the constant-ratio characterization of the tradeoff curve into the constant-ratio result on the average cost.

Problem K (Decentralized LQG problem with average power constraints)**.** *Consider the same dynamics as Problem J. But, now the control objective is minimizing the state distortion for given input power constraints* $P_1, P_2 \in \mathbb{R}^+$*. We will say the power-distortion tradeoff,* $D(P_1, P_2)$ *is achievable if and only if there exist causal control strategies* $u_1[n], u_2[n]$ *such that*

$$\limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mathbb{E}[x^2[n]] \leq D(P_1, P_2),$$
$$\limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mathbb{E}[u_1^2[n]] \leq P_1,$$
$$\limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mathbb{E}[u_2^2[n]] \leq P_2.$$

5.3 Qualitative Understanding of Centralized LQG Problems

Before we present the technical details, we first explain the insight behind the results. Theorem 5.1 states that control by a single controller is enough to achieve an approximately optimal performance. The optimal control by a single controller is a well-studied LQG control problem. The optimal average cost, the weighted sum of the input power and the state distortion, is the solution of a Riccati equation.

However, even though Riccati equations give exact optimal costs for centralized control problems, their quantitative results are hard to interpret. Therefore, in this section, we will approximate the optimal costs to simple functions, so that we can gain intuitive and qualitative understanding about the centralized control problems. Furthermore, we will take take a distortion-power-tradeoff perspective rather than a minimum-cost point-of-view.

Let's first formally state the scalar centralized LQG problems.

Problem L (Centralized LQG with average power constraints)**.** *Consider the following dynamic system with a single controller.*

$$x[n+1] = ax[n] + u[n] + w[n]$$
$$y[n] = x[n] + v[n]$$

where $x[0] \sim \mathcal{N}(0, \sigma_0^2)$*,* $w[n] \sim \mathcal{N}(0, 1)$*,* $v[n] \sim \mathcal{N}(0, \sigma_v^2)$ *are independent Gaussian random variables. The control input* $u[n]$ *must be a causal function of* $y[n]$*, i.e.* $u[n] = f_n(y_1[0], \cdots, y_1[n])$*.*

The control objective is minimizing the state distortion for a given input power constraint $P \in \mathbb{R}^+$*. We say the power-distortion tradeoff* $D_{\sigma_v}(P)$ *is achievable if and only if there exists a* *causal control strategy* $u[n]$ *such that*

$$\limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mathbb{E}[x^2[n]] \le D_{\sigma_v}(P),$$
$$\limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mathbb{E}[u^2[n]] \le P.$$

Definition 5.2 (Optimal Linear Strategy $L_{lin,cen}$ for Centralized LQG problems)**.** *Consider the centralized LQG problem of Problem L. Let* $L_{lin,cen}$ *be the set of all controllers which can be written in the following form. For some* $k \in \mathbb{R}$, $u[n] = -k\mathbb{E}[x[n]|y^n, u^n]$.

Lemma 5.1. *Consider the centralized LQG problem of Problem L. Define* Σ_E *as*

$$\Sigma_E := \frac{(a^2-1)\sigma_v^2 - 1 + \sqrt{((a^2-1)\sigma_v^2-1)^2 + 4a^2\sigma_v^2}}{2a^2}. \tag{5.2}$$

Then, for all k *such that* $|a-k| < 1$, *the linear strategy of Definition 5.2 can achieve the following Power-distortion tradeoff:*

$$(D_{\sigma_v}(P), P) = (\frac{(2ak-k^2)\Sigma_E + 1}{1-(a-k)^2}, k^2(\frac{(2ak-k^2)\Sigma_E+1}{1-(a-k)^2} - \Sigma_E)). \tag{5.3}$$

Furthermore, this power-distortion tradeoff is optimal in the sense that for a given P, *there is no control strategy which can achieve an expected squared state smaller than* $D_{\sigma_v}(P)$.

Proof. Let $\hat{x}[n] := \mathbb{E}[x[n]|y^n, u^{n-1}]$. Since $u[n] = -k\hat{x}[n]$,

$$\begin{aligned} x[n+1] &= ax[n] - k\hat{x}[n] + w[n] \\ &= a(x[n]-\hat{x}[n]) + (a-k)\hat{x}[n] + w[n]. \end{aligned} \tag{5.4}$$

Define $\Sigma_{X,n} := \mathbb{E}[x^2[n]]$, $\Sigma_{\hat{X},n} := \mathbb{E}[\hat{x}^2[n]]$, $\Sigma_{E,n} := \mathbb{E}[(x[n]-\hat{x}[n])^2]$. Then, we have

$$\begin{aligned} \Sigma_{X,n} &= \mathbb{E}[(x[n]-\hat{x}[n]+\hat{x}[n])^2] \\ &= \mathbb{E}[(x[n]-\hat{x}[n])^2] + \mathbb{E}[\hat{x}^2[n]] \\ &= \Sigma_{E,n} + \Sigma_{\hat{X},n} \end{aligned} \tag{5.5}$$

where the second equality comes from the orthogonality of $x[n]-\hat{x}[n]$ and $\hat{x}[n]$. Likewise, by (5.4) we also have

$$\begin{aligned} \Sigma_{X,n+1} &= a^2\Sigma_{E,n} + (a-k)^2\Sigma_{\hat{X},n} + 1 \\ &= a^2\Sigma_{E,n} + (a-k)^2(\Sigma_{X,n} - \Sigma_{E,n}) + 1 \end{aligned} \tag{5.6}$$

where the last inequality comes from (5.5).

Moreover, it is well-known that Kalman filtering performance converges to a steady state. In other words, by [11] we have

$$\Sigma_E := \lim_{n\to\infty} \Sigma_{E,n} = \frac{(a^2-1)\sigma_v^2 - 1 + \sqrt{((a^2-1)\sigma_v^2-1)^2 + 4a^2\sigma_v^2}}{2a^2}$$

Thus, by (5.6), $\Sigma_{X,n}$ converges as long as $|a-k|<1$. Let $\lim_{n\to\infty}\Sigma_{X,n} = \Sigma_X$. Then, by (5.6) we have

$$\begin{aligned}\Sigma_X &= \frac{(a^2-(a-k)^2)\Sigma_E + 1}{1-(a-k)^2} \\ &= \frac{(2ak-k^2)\Sigma_E+1}{1-(a-k)^2}.\end{aligned}$$

Since $u[n] = -k\hat{x}[n]$, using (5.5) the input power converges as follows.

$$\begin{aligned}\lim_{n\to\infty} \mathbb{E}[u^2[n]] &= k^2(\Sigma_X - \Sigma_E) \\ &= k^2(\frac{(2ak-k^2)\Sigma_E+1}{1-(a-k)^2} - \Sigma_E).\end{aligned}$$

This finishes the achievability proof of the tradeoff. The tightness of the tradeoff and the optimality of centralized linear controllers are well-known in the community, and we refer to [11] for a rigorous proof based on dynamic programming. □

As mentioned in the proof, Σ_E represents the Kalman filtering performance (mean square estimation error).

In the following discussions, we will qualitatively understand the tradeoff between the state distortion and control power by dividing into cases based on the eigenvalue of the system.

5.3.1 When $|a| = 1$

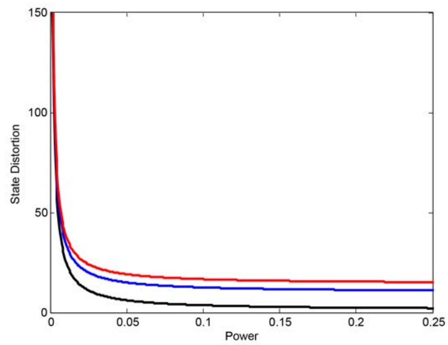


Figure 5.1: The Optimal State Distortion-Input Power Tradeoff: When $a=1$ with different values of σ_v^2 ($\sigma_v^2 = 1$ (Black line), $\sigma_v^2 = 100$ (Blue line), $\sigma_v^2 = 200$ (Red line))

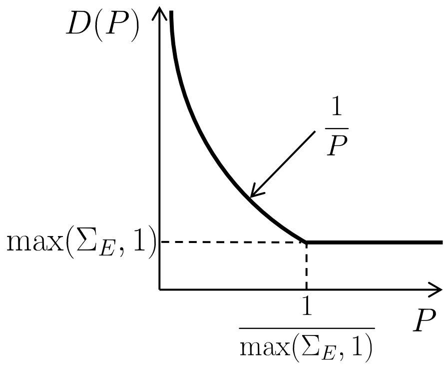


Figure 5.2: Conceptual Plot of State Distortion-Input Power Tradeoff: When $|a| = 1$

First, let's consider the case when the magnitude of the eigenvalue is 1, i.e. $|a| = 1$.

Since the Kalman filtering performance Σ_E is the minimum squared error for estimating the states, we can see $D_{\sigma_v}(P) \geq a^2\Sigma_E + 1$ for all P. For notational convenience, let's approximate $a^2\Sigma_E + 1$ by $\max(\Sigma_E, 1)$.

To achieve $D_{\sigma_v}(P) \approx \max(\Sigma_E, 1)$, the control power P has to be large enough. As we can see in Figure 5.1, the state distortion $D_{\sigma_v}(P)$ inversely proportionally increases as the control power P decreases.

Therefore, the power-state distortion tradeoff $D_{\sigma_v}(P)$ can be conceptualized as Figure 5.2. When the power P is smaller than $\frac{1}{\max(\Sigma_E,1)}$, the state distortion behaves like $\frac{1}{P}$. When the power becomes greater than $\frac{1}{\max(\Sigma_E,1)}$, the state distortion saturates at $\max(\Sigma_E, 1)$.

Let's write $(a_1, \cdots, a_n) \geq (b_1, \cdots, b_n)$ if and only if $a_1 \geq b_1, \cdots, a_n \geq b_n$. Then, Corollary 5.1 shows the formal statement of the power-distortion tradeoff for the centralized LQG problem.

Corollary 5.1. *Consider the centralized LQG problem shown in Problem L. When $|a| = 1$, the achievable power-distortion tradeoff $(D_{\sigma_v}(P), P)$ by the strategies of Definition 5.2 is upper bounded as follows:*

$$(D_{\sigma_v}(P), P) \leq (\frac{2}{t}, t) \text{ for all } 0 < t \leq \frac{1}{\max(1, \Sigma_E)} \quad (5.7)$$

where the definition of Σ_E is given as (5.2).

Especially, when $\sigma_v \geq 16$, we have

$$(D_{\sigma_v}(P), P) \leq (\frac{2}{t}, t) \text{ for all } 0 < t \leq \frac{1}{1.0005\sigma_v}. \quad (5.8)$$

When $\sigma_v \leq 16$, we have

$$(D_{\sigma_v}(P), P) \leq (\frac{2}{t}, t) \text{ for all } 0 < t \leq \frac{1}{15.008}. \quad (5.9)$$

Proof. See Appendix 10.1 for the proof. □

As we can see from (5.7), for powers $0 < P \leq \frac{1}{\max(1,\Sigma_E)}$, the tradeoff is inversely proportional. When the power becomes $P = \frac{1}{\max(1,\Sigma_e)}$, the state distortion saturates at the Kalman filtering performance.

In fact, careful inspection of Figure 5.1 shows that the transition between the interval $P \in [0, \frac{1}{\max(\Sigma_E,1)}]$ and $P \in [\frac{1}{\max(\Sigma_E,1)}, \infty]$ is much smoother than the one suggested in the conceptual plot of Figure 5.2. Therefore, a better approximation of the tradeoff can be $D_{\sigma_v}(P) \approx \frac{1}{P} + \max(\Sigma_E, 1)$ rather than $D_{\sigma_v}(P) \approx \max(\frac{1}{P}, \Sigma_E, 1)$ suggested in Figure 5.2. In fact, since $\max(\frac{1}{P}, \Sigma_E, 1) \leq \frac{1}{P} + \max(\Sigma_E, 1) \leq 2\max(\frac{1}{P}, \Sigma_E, 1)$, the two approximations are within a constant ratio. Thus, both approximations are enough to prove constant ratio optimality. In this chapter, we choose the approximation shown in Figure 5.2, since it is more discrete and thereby easier to compare with the lower bound in Section 5.4.2 by dividing cases.

Furthermore, we can prove that the optimal tradeoff $(D_{\sigma_v}(P), P)$ can be upper and lower bounded by the approximation of Figure 5.2 within a constant ratio. Consider the case[1] when $\sigma_v \geq 16$, then (5.8) of Lemma 5.1 gives an achievable upper bound on the tradeoff, $D_{\sigma_v}(P) \leq \frac{2}{P}$ for all $0 < P \leq \frac{1}{1.0005\sigma_v} \leq \frac{1}{16}$. We will see that Corollary 5.5 of Section 5.4.2 gives a lower bound on the tradeoff. By putting the second controller's noise $\sigma_{v2} = \infty$ and considering the first controller as the centralized controller, (b) of Corollary 5.5 gives that $D_{\sigma_v}(P) \geq \frac{0.02417}{P} + 1$ for all $P \leq \frac{1}{64}$. Therefore, we can notice that the upper and lower bound match within a constant ratio. Moreover, (d) of Corollary 5.5 gives that $D_{\sigma_v}(P) \geq \max(\frac{\sqrt{2}}{2}\sigma_{v1}, 1)$ for all P, which justifies the flat part of Figure 5.2. Therefore, increasing input power P more than $\frac{1}{1.0005\sigma_v}$ will not be greatly helpful, and we can use an achievable upper bound $D_{\sigma_v}(P) = 2.001\sigma_v$ for all $P \geq \frac{1}{1.0005\sigma_v}$ to prove constant-ratio optimality. This constant-ratio characterization of the tradeoff curve can be easily converted to a constant ratio optimality of average-cost problems by applying Lemma 4.14 of Chapter 4.

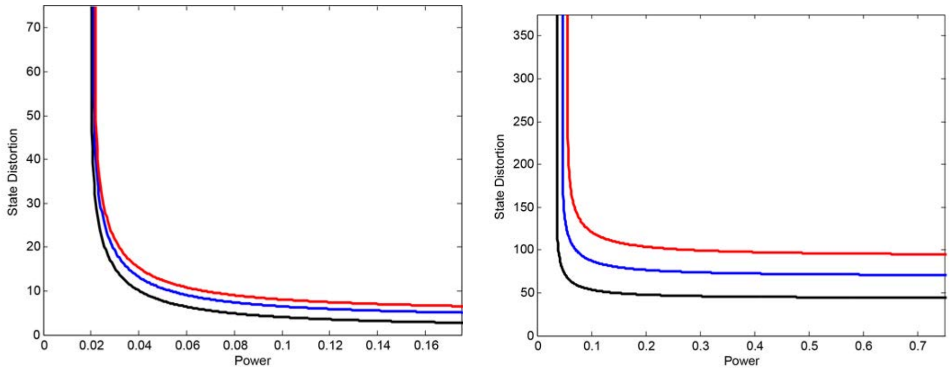
5.3.2 When $1 < |a| \leq 2.5$

Let's consider the case[2] when $1 < |a| \leq 2.5$. Just like the case of $|a| = 1$, the state distortion saturates at $a^2\Sigma_E + 1 \approx \max(\Sigma_E, 1)$ for all P, and the state distortion inversely proportionally increases as the power decreases.

However, there is a significant difference from the previous case of $|a| = 1$. Since the system is unstable by itself, when the power is too small the state distortion diverges to infinity. Figure 5.3a shows this behavior. Furthermore, it is well known that the minimum capacity to stabilize unstable plants is $\log|a|$. Since the variance of $w[n]$ is 1, the capacity from the controller to the plant can be thought as of $\frac{1}{2}\log(1+P)$. Therefore, the stabilizability condition $\frac{1}{2}\log(1+P) > \log|a|$ gives $P \geq a^2 - 1$ to stabilize the system.

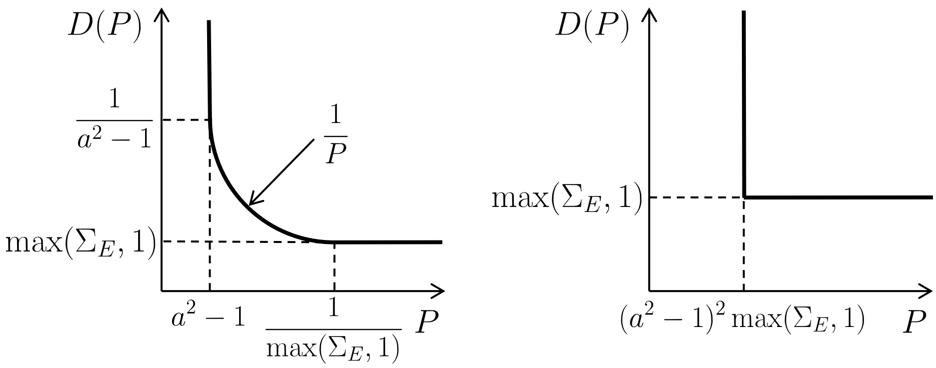
[1]Recall that when $|a| = 1$, $\Sigma_E \approx \sigma_v$.

[2]Here, the explicit number 2.5 does not have to be 2.5. In fact, we can choose any fixed number like $2, 3, 5, 6, \cdots$.



(a) (b)

Figure 5.3: The Optimal State Distortion-Input Power Tradeoff: $a = 1.01$. In (a), σ_v^2 are chosen as $\sigma_v^2 = 1$ (Black line), $\sigma_v^2 = 10$ (Blue line), $\sigma_v^2 = 20$ (Red line). In (b), σ_v^2 are chosen as $\sigma_v^2 = 1000$ (Black line), $\sigma_v^2 = 2000$ (Blue line), $\sigma_v^2 = 3000$ (Red line).



(a) (b)

Figure 5.4: Conceptual Plot of State Distortion-Input Power Tradeoff for $1 < |a| \le 2.5$: (a) is when $\max(\Sigma_E, 1) \le \frac{1}{a^2-1}$. (b) is when $\max(\Sigma_E, 1) \ge \frac{1}{a^2-1}$.

Based on the above discussion, we can draw a conceptual power-distortion tradeoff curve as shown in Figure 5.4a. Like Figure 5.2, when the power is larger than $\frac{1}{\max(\Sigma_E,1)}$, the state distortion saturates at $\max(\Sigma_E, 1)$. When the power is between $a^2 - 1$ and $\frac{1}{\max(\Sigma_E,1)}$, the state distortion is inversely proportional to the power. However unlike Figure 5.2 when the power is smaller than $(a^2 - 1)$, the controller cannot stabilize the system, so the state distortion diverges to infinity.

Furthermore, Figure 5.3b shows that as Σ_E increases, the gap between $(a^2 - 1)$ and $\frac{1}{\max(\Sigma_E,1)}$ (the interval where the distortion is inversely proportional to the power) decreases, i.e. the boundary of the optimal tradeoff region shrinks. Eventually, the whole boundary will converge to one point. Figure 5.4b conceptualizes this situation. When Σ_E is large enough so that $\max(\Sigma_E, 1) \geq \frac{1}{a^2-1}$, we need at least $(a^2 - 1)^2 \max(\Sigma_E, 1)$ controller power to stabilize the plant, and the corresponding state distortion saturates at the Kalman filtering performance $\max(\Sigma_E, 1)$.

The following corollary shows a formal statement of these conceptual tradeoff curves shown in Figure 5.4.

Corollary 5.2. *Consider the centralized single-controller LQG problem shown in Problem L. When $|a| > 1$, the achievable power-distortion tradeoff $(D_{\sigma_v}(P), P)$ by the strategies of Definition 5.2 is upper bounded as follows:*
(i) $(D_{\sigma_v}(P), P) \leq ((a^2 + 1)\Sigma_E + \frac{a^2}{a^2-1}, (a^2 - 1)^2\Sigma_E + (a^2 - 1))$
(ii) $(D_{\sigma_v}(P), P) \leq (\frac{4(|a|+1)^2}{t}, t)$ *for all* $2(|a| + 1)^2(1 - (\frac{1}{a})^2) \leq t \leq \frac{2(|a|+1)^2}{\max(1,(a^2+1)\Sigma_E)}$
where the definition of Σ_E *is given in* (5.2).

Especially, when $1 < |a| \leq 2.5$, $D_{\sigma_v}(P)$ *satisfies the following conditions:*
(i') $(D_{\sigma_v}(P), P) \leq (7.25\Sigma_E + \frac{6.25}{a^2-1}, (a^2 - 1)^2\Sigma_E + (a^2 - 1))$
(ii') $(D_{\sigma_v}(P), P) \leq (\frac{49}{t}, t)$ *for all* $8(a^2 - 1) \leq t \leq \frac{8}{\max(1,7.25\Sigma_E)}$

Proof. See Appendix 10.1 for the proof. □

When $\max(\Sigma_E, 1) \leq \frac{1}{a^2-1}$, (ii') of the corollary shows that we can achieve the tradeoff curve shown in Figure 5.4a. More precisely, when $P = 8(a^2 - 1)$, the statement (ii') reduces to $D_{\sigma_v}(P) \leq \frac{49}{8(a^2-1)}$. Therefore, $(D_{\sigma_v}(P), P) \approx (\frac{1}{a^2-1}, a^2 - 1)$ is achievable (up to constant scaling).

When $P = \frac{8}{\max(1,7.25\Sigma_E)}$, the statement (ii') reduces to $D_{\sigma_v}(P) \leq \frac{49}{8}\max(1, 7.25\Sigma_E)$. Thus, $(D_{\sigma_v}(P), P) \approx (\max(\Sigma_E, 1), \frac{1}{\max(\Sigma_E,1)})$ is also achievable. Between these two values, the tradeoff is inversely proportional.

When $\max(\Sigma_E, 1) \geq \frac{1}{a^2-1}$, (i') of the corollary shows the tradeoff curve in Figure 5.4b is achievable. More precisely, with the condition $\max(\Sigma_E, 1) \geq \frac{1}{a^2-1}$, the statement (i') implies

$$
\begin{aligned}
(D_{\sigma_v}(P), P) &\leq (7.25\Sigma_E + \frac{6.25}{a^2 - 1}, (a^2 - 1)^2\Sigma_E + (a^2 - 1)) \\
&\leq (13.5 \max(\Sigma_E, 1), 2(a^2 - 1)^2 \max(\Sigma_E, 1)).
\end{aligned}
$$

Therefore, the corner point of Figure 5.4b is achievable up to constant scaling. The whole tradeoff region is also achievable since we can always achieve the points with more state distortion and input power.

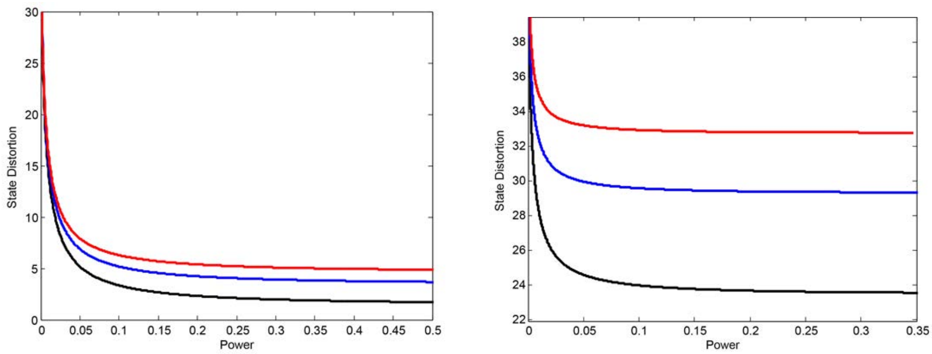
Just like Section 5.3.1, a careful inspection of Figure 5.3a suggests that $D_{\sigma_v}(P) \approx \frac{1}{P-(a^2-1)} + \max(\Sigma_E, 1)$ may be a better approximation than the one shown in Figure 5.4a. However, just like the discussion in Section 5.3.1, the approximation of Figure 5.3a is good enough to prove constant-ratio optimality, and easier to compare with a lower bound on the performance since the approximation is divided into multiple regions.

In fact, by putting $\Sigma_2 = \infty$ and considering the first controller as the centralized controller, (g), (f), (j) of Corollary 5.4 in Section 5.4.1 respectively reduce to

$$
\begin{aligned}
&D_{\sigma_v}(P) = \infty \text{ for all } P \leq \frac{1}{20}(a^2-1) \\
&D_{\sigma_v}(P) \geq \frac{0.0006976}{P_1} + 1 \text{ for all } P \leq \frac{1}{150} \\
&D_{\sigma_v}(P) \geq \max(0.1035\Sigma_1, 1).
\end{aligned}
$$

By taking the maximum over these three bounds, we can easily check that the resulting lower bound coincides with the approximation of Figure 5.3a up to a constant, and thereby the average cost can also be characterized within a constant by Lemma 4.14 of Chapter 4.

5.3.3 When $0.9 \leq |a| < 1$



(a) (b)

Figure 5.5: The Optimal State Distortion-Input Power Tradeoff: $a = 0.99$. In (a), σ_v^2 are chosen as $\sigma_v^2 = 1$ (Black line), $\sigma_v^2 = 10$ (Blue line), $\sigma_v^2 = 20$ (Red line). In (b), σ_v^2 are chosen as $\sigma_v^2 = 1000$ (Black line), $\sigma_v^2 = 2000$ (Blue line), $\sigma_v^2 = 3000$ (Red line).

Let's consider the case when $0.9 \leq |a| < 1$. In contrast to the case of $1 < |a| \leq 2.5$, the system is stable by itself in this case. Therefore, the state distortion never increases above $\frac{1}{1-a^2}$.

As we can see in Figure 5.5a, the essential tradeoff curve is similar to the case of $|a| = 1$. For all control powers P, the state distortion saturates at the Kalman filtering performance $\max(\Sigma_E, 1)$. For control powers between $1 - a^2$ and $\frac{1}{\max(\Sigma_E, 1)}$, the state distortion is inversely proportional to the control power.

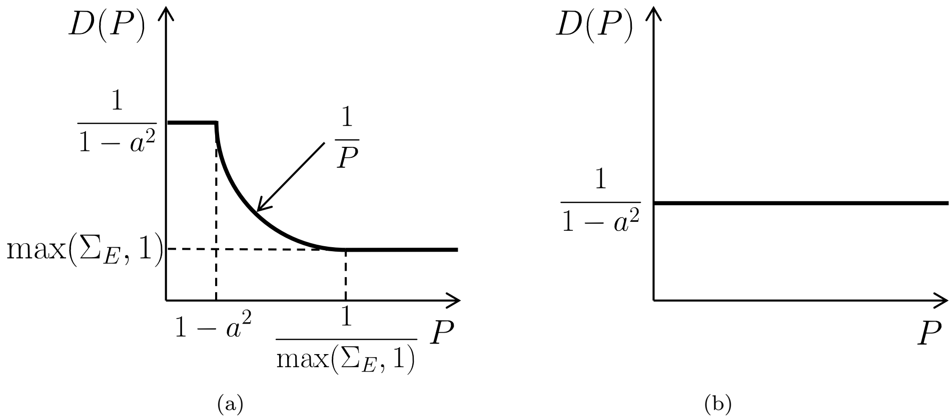


Figure 5.6: Conceptual Plot of State Distortion-Input Power Tradeoff for $0.9 \leq |a| < 1$: (a) is when $\max(\Sigma_E, 1) \leq \frac{1}{1-a^2}$. (b) is when $\max(\Sigma_E, 1) \geq \frac{1}{1-a^2}$.

However, when the power becomes smaller than $1 - a^2$, this inversely-proportional caricature of the state distortion becomes larger than $\frac{1}{1-a^2}$, which is achievable even without any control. Therefore, for powers smaller than $1 - a^2$, the state distortion stays at $\frac{1}{1-a^2}$. Therefore, the conceptual tradeoff curve has to follow the curve in Figure 5.4a.

Furthermore, Figure 5.5b shows that as Σ_E increases, the state distortion without control ($\frac{1}{1-a^2}$) and the Kalman filtering performance ($\max(\Sigma_E, 1)$) becomes similar. Eventually, when $\max(\Sigma_E, 1) \geq \frac{1}{1-a^2}$, as depicted in Figure 5.4b, the minimum state distortion becomes $\frac{1}{1-a^2}$ which is achievable even without any control.

Corollary 5.3 gives formal statements of these observations.

Corollary 5.3. *Consider the centralized LQG problem shown in Problem L. When $|a| < 1$, the achievable power-distortion tradeoff $(D_{\sigma_v}(P), P)$ by the strategies of Definition 5.2 is upper bounded as follows:*

$$(D_{\sigma_v}(P), P) \leq (\frac{1}{1-a^2}, 0), \tag{5.10}$$

and especially when[3] $\Sigma_E \leq \frac{1}{1-a^2}$ *we also have*

$$(D_{\sigma_v}(P), P) \leq (\frac{2}{t}, t) \text{ for all } 1 - a^2 \leq t \leq \frac{1}{\max(1, \Sigma_E)} \tag{5.11}$$

where the definition of Σ_E is given as (5.2).

Proof. See Appendix 10.1 for the proof. □

When $\max(\Sigma_E, 1) \geq \frac{1}{1-a^2}$, (5.10) shows the tradeoff curve shown in Figure 5.6b is achievable.

[3] Since $|a| < 1$, the condition $\Sigma_E \leq \frac{1}{1-a^2}$ is equivalent to the condition $\max(1, \Sigma_E) \leq \frac{1}{1-a^2}$.

When $\max(\Sigma_E, 1) \leq \frac{1}{1-a^2}$, (5.11) shows the inversely proportional tradeoff curve shown in Figure 5.6a when the power is between $1 - a^2$ and $\frac{1}{\max(1,\Sigma_E)}$.

In fact, in Figure 5.5a we cannot find a flat region for the power between 0 and $1 - a^2$ which is shown in the approximation of Figure 5.6a. Therefore, like Section 5.3.1, 5.3.2, a better approximation of the tradeoff might be $D_{\sigma_v}(P) \approx \frac{1}{P-(a^2-1)} + \max(\Sigma_E, 1)$ and worth exploring. However, the approximation of Figure 5.5a is good enough to give a constant-ratio optimality result. For example, if we compute the distortion for $P \in [0, 1 - a^2]$ with this new approximation, we get $D_{\sigma_v}(P) \in [\frac{1}{2(1-a^2)} + \max(\Sigma_E, 1), \frac{1}{1-a^2} + \max(\Sigma_E, 1)]$. Especially, for $\max(\Sigma_E, 1) \leq \frac{1}{1-a^2}$ which is the case of Figure 5.6a, this interval is included in

$$[\frac{1}{2(1-a^2)} + \max(\Sigma_E, 1), \frac{1}{1-a^2} + \max(\Sigma_E, 1)] \subseteq [\frac{1}{2(1-a^2)}, \frac{2}{1-a^2}].$$

Therefore, the approximation is essentially the same as the one of Figure 5.6a, $\frac{1}{1-a^2}$, up to a constant.

Furthermore, Corollary 5.6 of Section 5.4.3 gives a matching lower bound to the approximation of Figure 5.6a. First notice that as σ_{v2} goes to infinity, the Kalman filtering performance Σ_2 converges to $\frac{1}{1-a^2}$ which is the expected squared-state of the stable system without any control. Thus, by putting $\Sigma_2 = \frac{1}{1-a^2}$, thinking of the second controller as the centralized controller, and considering the case of $\frac{1}{1-a^2} \geq 40$, the conditions (a), (b), (e) of Corollary 5.6 respectively reduce to

$$\begin{aligned}
D_{\sigma_v}(P) &\geq \frac{0.009131}{1-a^2} + 1 \text{ for all } P \leq 1 - a^2 \\
D_{\sigma_v}(P) &\geq \frac{0.009131}{P} + 1 \text{ for all } 1 - a^2 \leq P \leq \frac{1}{40} \\
D_{\sigma_v}(P) &\geq \max(0.2636\Sigma_1, 1) \text{ for all } P.
\end{aligned}$$

Therefore, we can easily observe that by taking the maximum of these bounds, we get a matching lower bound to Figure 5.6a (up to a constant). Therefore, by Lemma 4.14 of Chapter 4, we can also characterize the average cost within a constant ratio using the approximation of Figure 5.6a.

5.3.4 When $|a| \leq 0.9$

In this case, the state distortion $\frac{1}{1-a^2}$ which can be obtained without any control input, is already small enough (smaller than 5.27). Therefore, the tradeoff curve is essentially the same as Figure 5.6b, which is achievable with zero control input.

5.4 Lower bounds and Constant-Ratio Results for Decentralized LQG problems

Now, we intuitively understand the power-distortion tradeoff of centralized single-controller LQG problems with scalar plants. Based on this understanding, we will prove that single-controller linear strategies are enough to achieve the optimal decentralized LQG performance to within a constant ratio. In other words, $(D(P_1, P_2), P_1, P_2)$ of Problem K is essentially $(\min(D_{\sigma_{v1}}(P_1), D_{\sigma_{v2}}(P_2)), P_1, P_2)$ where the definition of $D_{\sigma_v}(P)$ is given in Problem L.

For the upper bound on the optimal cost of the decentralized LQG problems, we can simply use the centralized controller's performance shown in Corollaries 5.2, 5.1, 5.3. However, we still need a lower bound on the cost of the decentralized LQG problems, and it turns out the naive lower bound we can obtain by merging two decentralized controllers into a centralized controller is too loose to prove constant-ratio optimality (One can easily see this by giving noiseless observation to one controller and zero input cost to the other.).

Therefore, in this section, we will give a non-trivial lower bound based on information theory [21] and prove that the proposed lower bounds are tight to within a constant ratio.

5.4.1 When $1 < |a| \leq 2.5$

The ideas for the lower bounds are essentially the same as the ones of Chapter 4. The main idea is geometric slicing, which can be thought of as a counterpart to cutset bounds in information theory [21]. The only difference from the geometric slicing lemma shown in Lemma 4.8 of Chapter 4 is that here we use allow arbitrary sequences for slicing the problem since we must use arithmetic sequences to slice the problem for the $|a| = 1$ case, and a geometric base that depends on a in the vicinity of $a = 1$.

As we did in Chapter 4, we first introduce sliced finite-horizon problems.

Problem M (Sliced Finite-horizon LQG problem for Problem J)**.** *Let the system equations, the problem parameters, the underlying random variables, and the restrictions on the controllers be given exactly the same as Problem J. However, now for given $k, k_1, k_2 \in \mathbb{N}(k_1 \leq k, k_2 \leq k)$ and positive sequences $\alpha_{k_1}, \alpha_{k_1+1}, \cdots, \alpha_{k-1}$ and $\beta_{k_2}, \beta_{k_1+1}, \cdots, \beta_{k-1}$, the control objective is*

$$\inf_{u_1,u_2} q\mathbb{E}[x^2[k]] + r_1 \sum_{k_1 \leq i \leq k-1} \alpha_i \mathbb{E}[u_1^2[i]] + r_2 \sum_{k_2 \leq i \leq k-1} \beta_i \mathbb{E}[u_2^2[n]].$$

Lemma 5.2 (Geometric Slicing)**.** *Let the system equations, the problem parameters, the underlying random variables, and the restrictions on the controllers be given as in Problem J. When $\sigma_0^2 = 0$, for all $k, k_1, k_2 \in \mathbb{N}(k_1 \leq k, k_2 \leq k)$ and positive sequences $\alpha_{k_1}, \alpha_{k_1+1}, \cdots, \alpha_{k-1}$ and $\beta_{k_2}, \beta_{k_1+1}, \cdots, \beta_{k-1}$ such that $\sum_{k_1 \leq i \leq k-1} \alpha_i = 1$ and $\sum_{k_2 \leq i \leq k-1} \beta_i = 1$, the infinite-horizon cost of Problem J is lower bounded by the finite-horizon cost of Problem M, i.e.*

$$\begin{aligned}
&\inf_{u_1,u_2} \limsup_{N \to \infty} \frac{1}{N} \sum_{0 \leq n \leq N-1} (q\mathbb{E}[x^2[n]] + r_1 \mathbb{E}[u_1^2[n]] + r_2 \mathbb{E}[u_2^2[n]]) \\
&\geq \inf_{u_1,u_2} q\mathbb{E}[x^2[k]] + r_1 \sum_{k_1 \leq i \leq k-1} \alpha_i \mathbb{E}[u_1^2[i]] + r_2 \sum_{k_2 \leq i \leq k-1} \beta_i \mathbb{E}[u_2^2[n]].
\end{aligned}$$

Furthermore, both costs are increasing functions of σ_0^2 *and when* $\sigma_0^2 = 0$, $u_1[0] = 0$ *and* $u_2[0] = 0$ *are optimal for both problems.*

Proof. The proof is essentially the same as the proof of Lemma 4.8 of Chapter 4. The only difference is that the geometric sequences used in Lemma 4.8 of Chapter 4 have to be replaced by α_n and β_n. $\square$

Using this lemma, we can lower bound on the cost of the decentralized LQG problems as follows. (Notice that the following lemma holds for all $|a| > 1$. However, it fails to give a constant ratio result for fast-dynamics case of $|a| \geq 2.5$ since it does not reflect the large deviation aspect of disturbance random variables.)

Lemma 5.3. *Define* S_L *as the set of* (k_1, k_2, k) *such that* $k_1, k_2, k \in \mathbb{N}$ *and* $1 \leq k_1 \leq k_2 \leq k$. *We also define* $D_{L,1}(\widetilde{P_1}, \widetilde{P_2}, k_1, k_2, k)$ *as follows:*

$$D_{L,1}(\widetilde{P_1}, \widetilde{P_2}; k_1, k_2, k) := (\sqrt{\frac{\Sigma + a^{2(k-k_1)}\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}}}{2^{2I'(\widetilde{P_1})}} + a^{2(k-k_2)}\frac{1-a^{-2(k-k_2)}}{1-a^{-2}}} \\ - \sqrt{a^{2(k-k_1-1)}\frac{(1-a^{-(k-k_1)})^2}{(1-a^{-1})^2}\widetilde{P_1}} - \sqrt{a^{2(k-k_2-1)}\frac{(1-a^{-(k-k_2)})^2}{(1-a^{-1})^2}\widetilde{P_2}})_+^2 + 1$$

where

$$\begin{aligned}
\Sigma &= \frac{a^{2(k-1)}\frac{1-a^{-2(k_1-1)}}{1-a^{-2}}}{2^{2I}} \\
I &= \frac{k_1-1}{2}\log(1+\frac{1}{(k_1-1)\sigma_{v1}^2}\frac{a^{2(k_1-2)}(1-a^{-2(k_1-1)})}{1-a^{-2}}\frac{1-a^{-2(k_1-1)}}{1-a^{-2}}) \\
&+ \frac{k_1-1}{2}\log(1+\frac{1}{(k_1-1)\sigma_{v2}^2}\frac{a^{2(k_1-2)}(1-a^{-2(k_1-1)})}{1-a^{-2}}\frac{1-a^{-2(k_1-1)}}{1-a^{-2}}) \\
I'(\widetilde{P_1}) &= \frac{k_2-k_1}{2}\log(1+\frac{1}{(k_2-k_1)\sigma_{v2}^2}(2a^{2(k_2-1-k)}\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}}\Sigma \\
&+ 2a^{2(k_2-1-k_1)}\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}}\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}} \\
&+ 2a^{2(k_2-k_1-2)}\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}}\frac{(1-a^{-(k_2-1-k_1)})(1-a^{-(k-k_1)})}{(1-a^{-1})^2}\widetilde{P_1})).
\end{aligned}$$

Here, when $k_1 - 1 = 0$, $I = 0$ *and when* $k_2 - k_1 = 0$, $I'(\widetilde{P_1}) = 0$.

Let $|a| > 1$. *Then, for all* $q, r_1, r_2, \sigma_0, \sigma_{v1}, \sigma_{v2} \geq 0$, *the minimum cost* (5.1) *of Problem J is lower bounded as follows:*

$$\begin{aligned}
&\inf_{u_1,u_2} \limsup_{N\to\infty} \frac{1}{N}\sum_{0\leq n<N} q\mathbb{E}[x^2[n]] + r_1\mathbb{E}[u_1^2[n]] + r_2\mathbb{E}[u_2^2[n]] \\
&\geq \sup_{(k_1,k_2,k)\in S_L} \min_{\widetilde{P_1},\widetilde{P_2}\geq 0} qD_{L,1}(\widetilde{P_1}, \widetilde{P_2}; k_1, k_2, k) + r_1\widetilde{P_1} + r_2\widetilde{P_2}.
\end{aligned}$$

Proof. For simplicity, we assume $a > 1$, $1 < k_1 < k_2 < k$. The remaining cases when $a < -1$ or $k_1 = 1$ or $k_2 - k_1 = 0$ or $k = k_2$ easily follow with minor modifications.

• Geometric Slicing: We apply the geometric slicing idea of Lemma 5.2 to get a finite-horizon problem. By putting $\alpha_{k_1} = (\frac{1-a^{-1}}{1-a^{-(k-k_1)}})$, $\alpha_{k_1+1} = (\frac{1-a^{-1}}{1-a^{-(k-k_1)}})a^{-1}, \cdots, \alpha_k = (\frac{1-a^{-1}}{1-a^{-(k-k_1)}})a^{-k+1+k_1}$ and $\beta_{k_2} = (\frac{1-a^{-1}}{1-a^{-(k-k_2)}})$, $\beta_{k_2+1} = (\frac{1-a^{-1}}{1-a^{-(k-k_2)}})a^{-1}$, $\cdots$, $\beta_{k-1} = (\frac{1-a^{-1}}{1-a^{-(k-k_2)}})a^{-k+1+k_2}$ the average cost is lower bounded by

$$\inf_{u_1,u_2}(q\mathbb{E}[x^2[k]]$$

$$+ r_1 \underbrace{((\frac{1-a^{-1}}{1-a^{-(k-k_1)}})\mathbb{E}[u_1^2[k_1]] + (\frac{1-a^{-1}}{1-a^{-(k-k_1)}})a^{-1}\mathbb{E}[u_1^2[k_1+1]] + \cdots + (\frac{1-a^{-1}}{1-a^{-(k-k_1)}})a^{-k+1+k_1}\mathbb{E}[u_1^2[k-1]])}_{:=\widetilde{P_1}}$$

$$+ r_2 \underbrace{((\frac{1-a^{-1}}{1-a^{-(k-k_2)}})\mathbb{E}[u_2^2[k_2]] + (\frac{1-a^{-1}}{1-a^{-(k-k_2)}})a^{-1}\mathbb{E}[u_2^2[k_2+1]] + \cdots + (\frac{1-a^{-1}}{1-a^{-(k-k_2)}})a^{-k+1+k_2}\mathbb{E}[u_2^2[k-1]]))}_{:=\widetilde{P_2}}$$

Here, we denote the second and third terms as $\widetilde{P_1}$ and $\widetilde{P_2}$ respectively.

• Three stage division: As we did Chapter 4, we will divide the finite-horizon problem into three time intervals — information-limited interval, Witsenhausen's interval, power-limited interval. Define

$$\begin{aligned}
W_1 &:= a^{k-1}w[0] + \cdots + a^{k-k_1+1}w[k_1-2] \\
W_2 &:= a^{k-k_1}w[k_1-1] + \cdots + a^{k-k_2+1}w[k_2-2] \\
W_3 &:= a^{k-k_2}w[k_2-1] + \cdots + aw[k-2] \\
U_{11} &:= a^{k-2}u_1[1] + \cdots + a^{k-k_1}u_1[k_1-1] \\
U_{21} &:= a^{k-2}u_2[1] + \cdots + a^{k-k_1}u_2[k_1-1] \\
U_1 &:= a^{k-k_1-1}u_1[k_1] + \cdots + u_1[k-1] \\
U_{22} &:= a^{k-k_1-1}u_2[k_1] + \cdots + a^{k-k_2}u_2[k_2-1]) \\
U_2 &:= a^{k-k_2-1}u_2[k_2] + \cdots + u_2[k-1] \\
X_1 &:= W_1 + U_{11} + U_{21} \\
X_2 &:= W_2 + U_{22}
\end{aligned}$$

W_1, W_2, W_3 represent the distortions of three intervals respectively. U_{11} and U_{21} respectively represent the first and second controller inputs in the information-limited interval. U_1 represents the remaining input of the first controller. U_{22} and U_2 represent the second controller's inputs in Witsenhausen's and power-limited intervals respectively.

The goal of this proof is grouping control inputs and expanding $x[n]$, so that we reveal the effects of the controller inputs on the state and isolate their effects according to their characteristics.

• Power-Limited Inputs: We will first isolate the power-limited inputs of the controllers, i.e. the first controller's input in the Witsenhausen's and power-limited intervals, and the second controller's input in the power-limited interval. Notice that

$$\begin{aligned}
x[k] &= w[k-1] + aw[k-2] + \cdots + a^{k-1}w[0] \\
&\quad + u_1[k-1] + au_1[k-2] + \cdots + a^{k-1}u_1[0] \\
&\quad + u_2[k-1] + au_2[k-2] + \cdots + a^{k-1}u_2[0] \\
&= (a^{k-1}w[0] + \cdots + a^{k-k_1+1}w[k_1-2] \\
&\quad + a^{k-2}u_1[1] + \cdots + a^{k-k_1}u_1[k_1-1] \\
&\quad + a^{k-2}u_2[1] + \cdots + a^{k-k_1}u_2[k_1-1]) \\
&\quad + (a^{k-k_1}w[k_1-1] + \cdots + a^{k-k_2+1}w[k_2-2] \\
&\quad + a^{k-k_1-1}u_2[k_1] + \cdots + a^{k-k_2}u_2[k_2-1]) \\
&\quad + (a^{k-k_2}w[k_2-1] + \cdots + aw[k-2]) \\
&\quad + (a^{k-k_1-1}u_1[k_1] + \cdots + u_1[k-1]) \\
&\quad + (a^{k-k_2-1}u_2[k_2] + \cdots + u_2[k-1]) \\
&\quad + w[k-1].
\end{aligned}$$

Therefore, by Lemma 4.1 of Chapter 4 we have

$$\begin{aligned}
\mathbb{E}[x^2[k]] &= \mathbb{E}[(X_1 + X_2 + W_3 + U_1 + U_2 + w[k-1])^2] \\
&= \mathbb{E}[(X_1 + X_2 + W_3 + U_1 + U_2)^2] + \mathbb{E}[w^2[k-1]] \\
&\geq (\sqrt{\mathbb{E}[(X_1 + X_2 + W_3)^2]} - \sqrt{\mathbb{E}[U_1^2]} - \sqrt{\mathbb{E}[U_2^2]})_+^2 + 1 \\
&= (\sqrt{\mathbb{E}[(X_1 + X_2)^2] + \mathbb{E}[W_3^2]} - \sqrt{\mathbb{E}[U_1^2]} - \sqrt{\mathbb{E}[U_2^2]})_+^2 + 1. \qquad (5.12)
\end{aligned}$$

where the last equality comes from causality. Here, we can see that $\mathbb{E}[(X_1 + X_2)^2]$ does not depend on the power-limited inputs.

• Information-Limited Interval: We will bound the remaining state distortion after the information-limited interval. Define y_1' and y_2' as follows.

$$\begin{aligned}
y_1'[k] &= a^{k-1}w[0] + a^{k-2}w[1] + \cdots + w[k-1] + v_1[k] \\
y_2'[k] &= a^{k-1}w[0] + a^{k-2}w[1] + \cdots + w[k-1] + v_2[k]
\end{aligned}$$

Here, $y_1'[k]$, $y_2'[k]$ can be obtained by removing $u_1[1:k-1]$, $u_2[1:k-1]$ from $y_1[k]$, $y_2[k]$, and $u_1[k]$ and $u_2[k]$ are functions of $y_1[1:k]$ and $y_2[1:k]$ respectively. Therefore, we can see that $y_1[1:k], y_2[1:k]$ are functions of $y_1'[1:k], y_2'[1:k]$. Moreover $W_1, y_1'[1:k_1-1], y_2'[1:k_1-1]$ are jointly Gaussian.

Let

$$W_1' := W_1 - \mathbb{E}[W_1|y_1'[1:k_1-1], y_2'[1:k_1-1]]$$
$$W_1'' := \mathbb{E}[W_1|y_1'[1:k_1-1], y_2'[1:k_1-1]].$$

Then, W_1', W_1'', W_2 are independent Gaussian random variables. Moreover, W_1', W_2 are independent from $y_1'[1:k_1-1], y_2'[1:k_1-1]$. W_1'' is a function of $y_1'[1:k_1-1], y_2'[1:k_1-1]$.

Now, let's lower bound $\mathbb{E}[(X_1+X_2)^2]$. Since Gaussians maximize entropy, we have

$$\begin{aligned}
&\frac{1}{2}\log(2\pi e\mathbb{E}[(X_1+X_2)^2] \\
&\geq h(X_1+X_2) \\
&\geq h(X_1+X_2|y_1'[1:k_1-1], y_2'[1:k_1-1], y_2[k_1:k_2-1]) \\
&= h(W_1'+W_1''+U_{11}+U_{12}+W_2+U_{22}|y_1'[1:k_1-1], y_2'[1:k_1-1], y_2[k_1:k_2-1]) \\
&= h(W_1'+W_2|y_1'[1:k_1-1], y_2'[1:k_1-1], y_2[k_1:k_2-1]). \qquad (5.13)
\end{aligned}$$

We will first lower bound the variance of W_1'. Notice that

$$\begin{aligned}
\mathbb{E}[y_1'[k]^2] &= a^{2(k-1)} + a^{2(k-2)} + \cdots + 1 + \sigma_{v1}^2 \\
&= a^{2(k-1)}\frac{1-a^{-2k}}{1-a^{-2}} + \sigma_{v1}^2
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}[y_2'[k]^2] &= a^{2(k-1)} + a^{2(k-2)} + \cdots + 1 + \sigma_{v2}^2 \\
&= a^{2(k-1)}\frac{1-a^{-2k}}{1-a^{-2}} + \sigma_{v2}^2.
\end{aligned}$$

Thus, we have

$$
\begin{aligned}
& I(W_1; y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
&= h(y_1'[1:k_1-1], y_2'[1:k_1-1]) - h(y_1'[1:k_1-1], y_2'[1:k_1-1]|W_1) \\
&\leq \sum_{1\leq i\leq k_1-1} h(y_1'[i]) + \sum_{1\leq i\leq k_1-1} h(y_2'[i]) - \sum_{1\leq i\leq k_1-1} h(v_1[i]) - \sum_{1\leq i\leq k_1-1} h(v_2[i]) \\
&\leq \sum_{1\leq k\leq k_1-1} \frac{1}{2}\log(\frac{a^{2(k-1)}\frac{1-a^{-2k}}{1-a^{-2}} + \sigma_{v1}^2}{\sigma_{v1}^2}) + \sum_{1\leq k\leq k_1-1} \frac{1}{2}\log(\frac{a^{2(k-1)}\frac{1-a^{-2k}}{1-a^{-2}} + \sigma_{v2}^2}{\sigma_{v2}^2}) \\
&= \frac{1}{2}\log(\prod_{1\leq k\leq k_1-1} \frac{a^{2(k-1)}\frac{1-a^{-2k}}{1-a^{-2}} + \sigma_{v1}^2}{\sigma_{v1}^2}) + \frac{1}{2}\log(\prod_{1\leq k\leq k_1-1} \frac{a^{2(k-1)}\frac{1-a^{-2k}}{1-a^{-2}} + \sigma_{v2}^2}{\sigma_{v2}^2}) \\
&\overset{(A)}{\leq} \frac{k_1-1}{2}\log(\frac{1}{k_1-1}\sum_{1\leq k\leq k_1-1} \frac{a^{2(k-1)}\frac{1-a^{-2k}}{1-a^{-2}} + \sigma_{v1}^2}{\sigma_{v1}^2}) + \frac{k_1-1}{2}\log(\frac{1}{k_1-1}\sum_{1\leq k\leq k_1-1} \frac{a^{2(k-1)}\frac{1-a^{-2k}}{1-a^{-2}} + \sigma_{v2}^2}{\sigma_{v2}^2}) \\
&= \frac{k_1-1}{2}\log(1 + \frac{1}{k_1-1}\sum_{1\leq k\leq k_1-1} \frac{a^{2(k-1)}\frac{1-a^{-2k}}{1-a^{-2}}}{\sigma_{v1}^2}) + \frac{k_1-1}{2}\log(1 + \frac{1}{k_1-1}\sum_{1\leq k\leq k_1-1} \frac{a^{2(k-1)}\frac{1-a^{-2k}}{1-a^{-2}}}{\sigma_{v2}^2}) \\
&\leq \frac{k_1-1}{2}\log(1 + \frac{1}{k_1-1}\sum_{1\leq k\leq k_1-1} \frac{a^{2(k-1)}\frac{1-a^{-2(k_1-1)}}{1-a^{-2}}}{\sigma_{v1}^2}) + \frac{k_1-1}{2}\log(1 + \frac{1}{k_1-1}\sum_{1\leq k\leq k_1-1} \frac{a^{2(k-1)}\frac{1-a^{-2(k_1-1)}}{1-a^{-2}}}{\sigma_{v2}^2}) \\
&= \frac{k_1-1}{2}\log(1 + \frac{1}{(k_1-1)\sigma_{v1}^2}\frac{a^{2(k_1-2)}(1-a^{-2(k_1-1)})}{1-a^{-2}}\frac{1-a^{-2(k_1-1)}}{1-a^{-2}}) \\
&+ \frac{k_1-1}{2}\log(1 + \frac{1}{(k_1-1)\sigma_{v2}^2}\frac{a^{2(k_1-2)}(1-a^{-2(k_1-1)})}{1-a^{-2}}\frac{1-a^{-2(k_1-1)}}{1-a^{-2}}) \qquad (5.14)
\end{aligned}
$$

(A): Arithmetic-Geometric mean.

Let's denote the R.H.S. of (5.14) as I. We also have

$$
\begin{aligned}
\mathbb{E}[W_1^2] &= a^{2(k-1)} + \cdots + a^{2(k-k_1+1)}) \\
&= a^{2(k-1)}(1 + \cdots + a^{-2(k_1-2)}) \\
&= a^{2(k-1)}\frac{1-a^{-2(k_1-1)}}{1-a^{-2}}. \qquad (5.15)
\end{aligned}
$$

Now, we can bound the variance of the Gaussian random variable W_1' as follows:

$$
\begin{aligned}
& \frac{1}{2}\log(2\pi e\mathbb{E}[W_1'^2]) = h(W_1') \\
& \geq h(W_1'|y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
& = h(W_1|y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
& = h(W_1) - I(W_1; y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
& \geq \frac{1}{2}\log(2\pi e a^{2(k-1)}\frac{1-a^{-2(k_1-1)}}{1-a^{-2}}) - I
\end{aligned}
$$

where the last inequality follows from (5.14) and (5.15).

Thus,

$$\mathbb{E}[W_1'^2] \geq \frac{a^{2(k-1)}\frac{1-a^{-2(k_1-1)}}{1-a^{-2}}}{2^{2I}} \tag{5.16}$$

and denote the R.H.S. of (5.16) as Σ. Since W_1' is Gaussian, we can write $W_1' = W_1''' + W_1''''$ where $W_1''' \sim \mathcal{N}(0, \Sigma)$, and W_1''', W_1'''' are independent.

Moreover, we also have

$$\begin{aligned}\mathbb{E}[W_2^2] &= a^{2(k-k_1)} + \cdots + a^{2(k-k_2+1)} \\ &= a^{2(k-k_1)}\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}}.\end{aligned} \tag{5.17}$$

By (5.13) we have

$$\begin{aligned}&\frac{1}{2}\log(2\pi e\mathbb{E}[(X_1+X_2)^2]) \\ &\geq h(W_1' + W_2 | y_1'[1:k_1-1], y_2'[1:k_1-1], y_2[k_1:k_2-1]) \\ &\geq h(W_1' + W_2 | W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1], y_2[k_1:k_2-1]) \\ &= h(W_1''' + W_2 | W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1], y_2[k_1:k_2-1]) \\ &= h(W_1''' + W_2 | W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\ &- I(W_1''' + W_2; y_2[k_1:k_2-1] | W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\ &= h(W_1''' + W_2) \\ &- I(W_1''' + W_2; y_2[k_1:k_2-1] | W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\ &\geq \frac{1}{2}\log(2\pi e(\Sigma + a^{2(k-k_1)}\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}})) \\ &- I(W_1''' + W_2; y_2[k_1:k_2-1] | W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1])\end{aligned} \tag{5.18}$$

where the last inequality follows from the fact that W_1''' and W_2 are independent Gaussian, and (5.16), (5.17).

Now, the question boils down to getting an upper bound of the last mutual information term, which can be understood as the information contained in the second controller's observation in Witsenhausen's interval.

• Second controller's observation in Witsenhausen's interval: We will bound the amount of information contained in the second controller's observation in Witsenhausen's interval. For $n \geq k_1$, define

$$\begin{aligned}y_2''[n] := &\, a^{n-k}W_1''' + a^{n-k_1}w[k_1-1] + a^{n-k_1-1}w[k_1] + \cdots + w[n-1] \\ &+ a^{n-k_1-1}u_1[k_1] + \cdots + u_1[n-1] \\ &+ v_2[n].\end{aligned}$$

Notice that the relationship between $y_2[n]$ and $y_2''[n]$ is

$$\begin{aligned}y_2[n] &= y_2''[n] + a^{n-k_1-1}u_2[k_1] + \cdots + u_2[n-1] \\ &\quad + a^{n-k}W_1'''' + a^{n-k}\mathbb{E}[W_1|y_1'[1:k_1-1], y_2'[1:k_1-1]]. \end{aligned} \tag{5.19}$$

The mutual information of (5.18) is bounded as follows:

$$\begin{aligned}
& I(W_1''' + W_2; y_2[k_1:k_2-1]|W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
& = h(y_2[k_1:k_2-1]|W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
& - h(y_2[k_1:k_2-1]|W_1''' + W_2, W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
& = \sum_{k_1 \le i \le k_2-1} h(y_2[i]|y_2[k_1:i-1], W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
& - \sum_{k_1 \le i \le k_2-1} h(y_2[i]|y_2[k_1:i-1], W_1''' + W_2, W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
& \overset{(A)}{=} \sum_{k_1 \le i \le k_2-1} h(y_2''[i]|y_2[k_1:i-1], W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
& - \sum_{k_1 \le i \le k_2-1} h(y_2[i]|y_2[k_1:i-1], W_1''' + W_2, W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
& \overset{(B)}{\le} \sum_{k_1 \le i \le k_2-1} h(y_2''[i]) - \sum_{k_1 \le i \le k_2-1} h(v_2[i]) \\
& \le \sum_{k_1 \le i \le k_2-1} \frac{1}{2}\log(2\pi e \mathbb{E}[y_2''[i]^2]) - \sum_{k_1 \le i \le k_2-1} \frac{1}{2}\log(2\pi e \sigma_{v2}^2)
\end{aligned} \tag{5.20}$$

(A): Since $y_2[1:k_1-1]$ is a function of $y_2'[1:k_1-1]$, $u_2[k_1], \cdots, u_2[i]$ are functions of $y_2[k_1:i-1]$, $y_2'[1:k_1-1]$. Thus, all the terms in (5.19) except $y_2''[i]$ vanish by the conditioning.
(B): By causality, $v_2[i]$ is independent from all conditioning random variables.

First, let's bound the variance of $y_2''[n]$. By Lemma 4.1 of Chapter 4, we have

$$\begin{aligned}
\mathbb{E}[y_2''[n]^2] &\le 2\mathbb{E}[(a^{n-k}W_1''' + a^{n-k_1}w[k_1-1] + a^{n-k_1-1}w[k_1] + \cdots + w[n-1])^2] \\
&\quad + 2\mathbb{E}[(a^{n-k_1-1}u_1[k_1] + \cdots + u_1[n-1])^2] + \sigma_{v2}^2 \\
&= 2(a^{2(n-k)}\Sigma + a^{2(n-k_1)} + \cdots + 1) \\
&\quad + 2\mathbb{E}[(a^{n-k_1-1}u_1[k_1] + \cdots + u_1[n-1])^2] + \sigma_{v2}^2.
\end{aligned}$$

Here, by putting $a = a$ and $b = a^{-1}$ to Lemma 4.10 of Chapter 4 we have

$$\begin{aligned}
& \mathbb{E}[(a^{n-k_1-1}u_1[k_1] + \cdots + u_1[n-1])^2] \\
& \le a^{2(n-k_1-1)}\frac{1-a^{-(n-k_1)}}{1-a^{-1}}(\mathbb{E}[u_1^2[k_1]] + a^{-1}\mathbb{E}[u_1^2[k_1+1]] + \cdots + a^{-(n-k_1-1)}\mathbb{E}[u_1^2[n-1]]) \\
& \le a^{2(n-k_1-1)}\frac{1-a^{-(n-k_1)}}{1-a^{-1}}\frac{1-a^{-(k-k_1)}}{1-a^{-1}}\widetilde{P_1} \\
& = a^{2(n-k_1-1)}\frac{(1-a^{-(n-k_1)})(1-a^{-(k-k_1)})}{(1-a^{-1})^2}\widetilde{P_1}.
\end{aligned}$$

Thus, the variance of $y_2''[n]$ is bounded as:

$$\mathbb{E}[y_2''[n]^2] \leq 2a^{2(n-k)}\Sigma + 2a^{2(n-k_1)}\frac{1-a^{-2(n-k_1+1)}}{1-a^{-2}} + 2a^{2(n-k_1-1)}\frac{(1-a^{-(n-k_1)})(1-a^{-(k-k_1)})}{(1-a^{-1})^2}\widetilde{P_1} + \sigma_{v2}^2.$$

Therefore, we have

$$\begin{aligned}
&\sum_{k_1 \leq n \leq k_2-1} \mathbb{E}[y_2''[n]^2] \\
&\leq \sum_{k_1 \leq n \leq k_2-1} 2a^{2(n-k)}\Sigma + 2a^{2(n-k_1)}\frac{1-a^{-2(n-k_1+1)}}{1-a^{-2}} + 2a^{2(n-k_1-1)}\frac{(1-a^{-(n-k_1)})(1-a^{-(k-k_1)})}{(1-a^{-1})^2}\widetilde{P_1} + \sigma_{v2}^2 \\
&\leq 2(a^{2(k_1-k)} + \cdots + a^{2(k_2-1-k)})\Sigma + \sum_{k_1 \leq n \leq k_2-1} 2a^{2(n-k_1)}\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}} \\
&+ \sum_{k_1 \leq n \leq k_2-1} 2a^{2(n-k_1-1)}\frac{(1-a^{-(k_2-1-k_1)})(1-a^{-(k-k_1)})}{(1-a^{-1})^2}\widetilde{P_1} + (k_2-k_1)\sigma_{v2}^2 \\
&\leq 2a^{2(k_2-1-k)}\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}}\Sigma + 2a^{2(k_2-1-k_1)}\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}}\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}} \\
&+ 2a^{2(k_2-k_1-2)}\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}}\frac{(1-a^{-(k_2-1-k_1)})(1-a^{-(k-k_1)})}{(1-a^{-1})^2}\widetilde{P_1} + (k_2-k_1)\sigma_{v2}^2 \qquad (5.21)
\end{aligned}$$

Therefore, by (5.20) and (5.21) we conclude

$$\begin{aligned}
&I(W_1''' + W_2; y_2[k_1 : k_2-1] | W'''', y_1'[1 : k_1-1], y_2'[1 : k_1-1]) \\
&\leq \sum_{k_1 \leq n \leq k_2-1} \frac{1}{2}\log(\frac{\mathbb{E}[y_2''[n]^2]}{\sigma_{v2}^2}) \\
&= \frac{1}{2}\log(\prod_{k_1 \leq n \leq k_2-1} \frac{\mathbb{E}[y_2''[n]^2]}{\sigma_{v2}^2}) \\
&\overset{(A)}{\leq} \frac{k_2-k_1}{2}\log(\frac{1}{k_2-k_1}\sum_{k_1 \leq n \leq k_2-1} \frac{\mathbb{E}[y_2''[n]^2]}{\sigma_{v2}^2}) \\
&\leq \frac{k_2-k_1}{2}\log(1 + \frac{1}{(k_2-k_1)\sigma_{v2}^2}(2a^{2(k_2-1-k)}\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}}\Sigma + 2a^{2(k_2-1-k_1)}\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}}\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}} \\
&+ 2a^{2(k_2-k_1-2)}\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}}\frac{(1-a^{-(k_2-1-k_1)})(1-a^{-(k-k_1)})}{(1-a^{-1})^2}\widetilde{P_1})) \qquad (5.22)
\end{aligned}$$

(A): Arithmetic-Geometric mean

Denote the R.H.S. of (5.22) as $I'(\widetilde{P_1})$. By (5.18), we can conclude

$$\frac{1}{2}\log(2\pi e\mathbb{E}[(X_1+X_2)^2] \geq \frac{1}{2}\log(2\pi e(\Sigma + a^{2(k-k_1)}\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}})) - I'(\widetilde{P_1})$$

which implies

$$\mathbb{E}[(X_1+X_2)^2] \geq \frac{\Sigma + a^{2(k-k_1)}\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}}}{2^{2I'(\widetilde{P_1})}}. \qquad (5.23)$$

• Final lower bound: Now, we can merge the inequalities to prove the lemma. The variance of W_3 is given as follows:

$$\mathbb{E}[W_3^2] = a^{2(k-k_2)} + \cdots + a^2 = a^{2(k-k_2)} \frac{1-a^{-2(k-k_2)}}{1-a^{-2}}. \tag{5.24}$$

By setting $a = a$ and $b = a^{-1}$ in Lemma 4.10 of Chapter 4, the variance of U_1 is bounded as follows:

$$\begin{aligned}\mathbb{E}[U_1^2] &\leq a^{2(k-k_1-1)} \frac{1-a^{-(k-k_1)}}{1-a^{-1}} (\mathbb{E}[u_1^2[k_1]] + a^{-1}\mathbb{E}[u_1^2[k_1+1]] + \cdots + a^{-(k-k_1-1)}\mathbb{E}[u_1^2[k-1]]) \\ &= a^{2(k-k_1-1)} \frac{(1-a^{-(k-k_1)})^2}{(1-a^{-1})^2} \widetilde{P_1}. \end{aligned} \tag{5.25}$$

Likewise, the variance of U_2 can be bounded as

$$\mathbb{E}[U_2^2] \leq a^{2(k-k_2-1)} \frac{(1-a^{-(k-k_2)})^2}{(1-a^{-1})^2} \widetilde{P_2}. \tag{5.26}$$

Finally, by plugging (5.23), (5.24), (5.25), (5.26) into (5.12), we prove the lemma. □

Corollary 5.4. *Consider the decentralized LQG problem of Problem J. Define*

$$\Sigma_1 := \frac{(a^2-1)\sigma_{v1}^2 - 1 + \sqrt{((a^2-1)\sigma_{v1}^2-1)^2 + 4a^2\sigma_{v1}^2}}{2a^2} \tag{5.27}$$

$$\Sigma_2 := \frac{(a^2-1)\sigma_{v2}^2 - 1 + \sqrt{((a^2-1)\sigma_{v2}^2-1)^2 + 4a^2\sigma_{v2}^2}}{2a^2}. \tag{5.28}$$

Let $1 < |a| \leq 2.5$. Then, for all $q, r_1, r_2, \sigma_0, \sigma_{v1}, \sigma_{v2} > 0$, the minimum cost (5.1) of Problem J is lower bounded as follows:

$$\begin{aligned}&\inf_{u_1,u_2} \limsup_{N\to\infty} \frac{1}{N} \sum_{0\leq n<N} q\mathbb{E}[x^2[n]] + r_1\mathbb{E}[u_1^2[n]] + r_2\mathbb{E}[u_2^2[n]] \\ &\geq \min_{\widetilde{P_1},\widetilde{P_2}\geq 0} qD_L(\widetilde{P_1},\widetilde{P_2}) + r_1\widetilde{P_1} + r_2\widetilde{P_2}\end{aligned}$$

where $D_L(\widetilde{P_1},\widetilde{P_2})$ satisfies the following conditions.

(a) If $\Sigma_1 \geq 150$, $\Sigma_2 \geq 150$, $\widetilde{P_1} \leq \frac{(a^2-1)^2\Sigma_1}{40000}$, $\widetilde{P_2} \leq \frac{(a^2-1)^2\Sigma_2}{40000}$ then $D_L(\widetilde{P_1},\widetilde{P_2}) = \infty$.
(b) If $\Sigma_1 \geq 150$, $\Sigma_2 \geq 150$, $\widetilde{P_1} \leq \frac{(a^2-1)^2\Sigma_1}{40000}$ then $D_L(\widetilde{P_1},\widetilde{P_2}) \geq 0.002774\Sigma_2 + 1$.
(c) If $\widetilde{P_1} \leq \frac{1}{20}(a^2-1)$, $\widetilde{P_2} \leq \frac{1}{20}(a^2-1)$ then $D_L(\widetilde{P_1},\widetilde{P_2}) = \infty$.
(d) If $\widetilde{P_1} \leq \frac{1}{75}$ and $\widetilde{P_2} \leq \frac{1}{75}$ then $D_L(\widetilde{P_1},\widetilde{P_2}) \geq 0.00389\frac{1}{\max(P_1,P_2)} + 1$.
(e) If $\Sigma_2 \geq 150$, $\widetilde{P_1} \leq \frac{1}{\Sigma_2}$ then $D_L(\widetilde{P_1},\widetilde{P_2}) \geq 0.0006976\Sigma_2 + 1$.
(f) If $\Sigma_2 \geq 150$, $\frac{1}{\Sigma_2} \leq \widetilde{P_1} \leq \frac{1}{150}$ then $D_L(\widetilde{P_1},\widetilde{P_2}) \geq \frac{0.0006976}{\widetilde{P_1}} + 1$.
(g) If $\Sigma_2 \geq 150$, $\widetilde{P_1} \leq \frac{1}{20}(a^2-1)$, $\widetilde{P_2} \leq \frac{(a^2-1)^2\Sigma_2}{40000}$ then $D_L(\widetilde{P_1},\widetilde{P_2}) = \infty$.
(h) If $\Sigma_1 \geq 150$, $\widetilde{P_1} \leq \frac{(a^2-1)^2\Sigma_1}{40000}$, $\widetilde{P_2} \leq \frac{1}{20}(a^2-1)$ then $D_L(\widetilde{P_1},\widetilde{P_2}) = \infty$.
(i) If $\Sigma_2 \geq 150$, $\widetilde{P_1} \leq \frac{1}{20}(a^2-1)$ then $D_L(\widetilde{P_1},\widetilde{P_2}) \geq 0.0002732\Sigma_2 + 1$.
(j) For all $\widetilde{P_1}$ and $\widetilde{P_2}$, $D_L(\widetilde{P_1},\widetilde{P_2}) \geq \max(0.1035\Sigma_1, 1)$.

Proof. See Appendix 10.2 for the proof. □

In this corollary, Σ_1 and Σ_2 are the Kalman filtering performance of the first and second controllers respectively.

Now, we have lower bounds on the average decentralized control cost of Problem J. Furthermore, by inspecting the form of the lower bounds, the term $D_L(\widetilde{P_1}, \widetilde{P_2})$ can be speculated as a lower bound on the power-distortion tradeoff $D(P_1, P_2)$ of Problem K.

Furthermore, Lemma 4.14 of Chapter 4 shows the average cost problem in Problem J and the power-distortion tradeoff problem in Problem K are closely related, i.e. if we can characterize the power-distortion tradeoff within a constant ratio, then we can characterize the average cost within a constant ratio. Therefore, in the following discussion, we will focus on the power-distortion tradeoff and justify that why it can be characterized within a constant. Throughout the discussion, we will consider $D_L(\widetilde{P_1}, \widetilde{P_2})$ as if it is a lower bound on $D(P_1, P_2)$ and the rigorous justification will be shown in Appendix sec:ageq1.

By comparing the achievable cost shown in Corollary 5.2, we will prove that they are within a constant ratio. In other words, we will prove the power-distortion tradeoff $(D(P_1, P_2), P_1, P_2)$ is essentially the better performance between two single controllers, i.e. $(\min(D_{\sigma_1}(P_1), D_{\sigma_2}(P_2)), P_1, P_2)$.

To justify this, we will divide the cases. As discussed in Section 5.3, the centralized controller's performance behaves qualitatively differently depending on $\max(\Sigma_1, 1)$, $\max(\Sigma_2, 1)$, $\frac{1}{a^2-1}$. Therefore, we will divide into three cases[4] depending on these values. Then, we will further divide the cases by P_1 and P_2.

When $\max(\Sigma_1, 1) \leq \max(\Sigma_2, 1) \leq \Theta(\frac{1}{a^2-1})$

We will again divide the cases based on P_1, P_2.

- When $P_1 \leq \Theta(a^2-1)$ and $P_2 \leq \Theta(a^2-1)$. As we can see from Figure 5.4a, each controller does not have enough power to stabilize the system. The statement (c) of Corollary 5.4 reveals that the system is unstable even in decentralized control problems.
- When $P_1 \leq \Theta(a^2-1)$ and $\Theta(a^2-1) \leq P_2 \leq \Theta(\frac{1}{\max(\Sigma_2,1)})$. In this case, the control performance is determined by the second controller. From Figure 5.4a we can see that $D(P_1, P_2) = O(\frac{1}{P_2})$ is achievable. The statement (d) of Corollary 5.4 tells it is tight up to a constant ratio.
- When $P_1 \leq \Theta(a^2-1)$ and $\Theta(\frac{1}{\max(\Sigma_2,1)}) \leq P_2$. Like above the second controller dominates the performance, and Figure 5.4a shows $D(P_1, P_2) = O(\max(\Sigma_2, 1))$. The statement (i) of Corollary 5.4 shows its tightness.
- When $\Theta(a^2-1) \leq P_1 \leq \Theta(\frac{1}{\max(\Sigma_2,1)})$ and $P_2 \leq \Theta(\frac{1}{\max(\Sigma_2,1)})$. In this case, the control performance is determined by the controller with larger power, and Figure 5.4a shows $D(P_1, P_2) = O(\frac{1}{\max(P_1,P_2)})$ is achievable. The statement (d) of Corollary 5.4 gives a matching lower bound.
- When $\Theta(a^2-1) \leq P_1 \leq \Theta(\frac{1}{\max(\Sigma_2,1)})$ and $\Theta(\frac{1}{\max(\Sigma_2,1)}) \leq P_2$. In this case, the second controller dominates the performance, and Figure 5.4a shows $D(P_1, P_2) = O(\Sigma_2, 1)$ is achievable.

[4]Since $\sigma_{v1} \leq \sigma_{v2}$, Σ_1 is always smaller than Σ_2.

The statement (e) of Corollary 5.4 gives a matching lower bound.

• When $\Theta(\frac{1}{\max(\Sigma_2,1)}) \leq P_1 \leq \Theta(\frac{1}{\max(\Sigma_1,1)})$. In this case, the first controller dominates the performance, and Figure 5.4a shows $D(P_1,P_2) = O(\frac{1}{P_1})$ is achievable. The statement (f) of Corollary 5.4 gives a matching lower bound.

• When $\Theta(\frac{1}{\max(\Sigma_1,1)}) \leq P_1$. In this case, the first controller dominates the performance, and Figure 5.4a shows $D(P_1,P_2) = O(\max(\Sigma_1,1))$ is achievable. The statement (j) of Corollary 5.4 gives a matching lower bound.

When $\max(\Sigma_1,1) \leq \Theta(\frac{1}{a^2-1}) \leq \max(\Sigma_2,1)$

We will further divide the cases based on P_1, P_2.

• When $P_1 \leq \Theta(a^2-1)$ and $P_2 \leq \Theta((a^2-1)^2\max(\Sigma_2,1))$. As we can see from Figure 5.4a, each controller does not have enough power to stabilize the system by itself. The statement (g) of Corollary 5.4 shows that the system is indeed necessarily unstable for decentralized control problems.

• When $P_1 \leq \Theta(a^2-1)$ and $\Theta((a^2-1)^2\max(\Sigma_2,1)) \leq P_2$. In this case, the second controller dominates the performance, and Figure 5.4b shows $D(P_1,P_2) = O(\max(\Sigma_2,1))$. The statement (i) of Corollary 5.4 give a matching lower bound up to a constant ratio.

• When $\Theta(a^2-1) \leq P_1 \leq \Theta(\frac{1}{\max(\Sigma_2,1)})$. Since we assume $\Theta(\frac{1}{a^2-1}) \leq \max(\Sigma_2,1)$, this case never happens.

• When $\Theta(\frac{1}{\max(\Sigma_2,1)}) \leq P_1 \leq \Theta(\frac{1}{\max(\Sigma_1,1)})$. In this case, the first controller dominates the performance, and Figure 5.4a shows $D(P_1,P_2) = O(\frac{1}{P_1})$ is achievable. The statement (f) of Corollary 5.4 gives a matching lower bound.

• When $\Theta(\frac{1}{\max(\Sigma_1,1)}) \leq P_1$. The first controller dominates the performance, but as we can see in Figure 5.4a its performance is saturated by Kalman filtering and $D(P_1,P_2) = O(\max(\Sigma_1,1))$. The statement (j) of Corollary 5.4 gives a matching lower bound.

When $\Theta(\frac{1}{a^2-1}) \leq \max(\Sigma_1,1) \leq \max(\Sigma_2,1)$

We will divide the cases based on P_1, P_2.

• When $P_1 \leq \Theta((a^2-1)^2\max(\Sigma_1,1))$ and $P_2 \leq \Theta((a^2-1)^2\max(\Sigma_1,1))$. In this case, as shown in Figure 5.4b each controller cannot stabilize the system by itself. The statement (a) of Corollary 5.4 shows that the decentralized system is indeed ncessarily unstable.

• When $P_1 \leq \Theta((a^2-1)^2\max(\Sigma_1,1))$ and $\Theta((a^2-1)^2\max(\Sigma_1,1)) \leq P_2$. In this case, the second controller dominates the performance, and Figure 5.4b shows $D(P_1,P_2) \leq O(\frac{1}{\max(\Sigma_2,1)})$ is achievable. The statement (b) of Corollary 5.4 gives a matching lower bound.

• When $\Theta((a^2-1)^2\max(\Sigma_1,1)) \leq P_1$. In this case, the first controller dominates the performance, and Figure 5.4b shows $D(P_1,P_2) \leq O(\max(\Sigma_1,1))$ is achievable. The statement (j) of Corollary 5.4 gives a matching lower bound.

Formally, the average cost can be characterized within a constant ratio as follows.

Proposition 5.1. *Consider the decentralized LQG control of Problem J. There exists $c \leq 2 \times 10^6$ such that for all $1 < |a| \leq 2.5$, q, r_1, r_2, σ_{v1} and σ_{v2},*

$$\frac{\inf_{u_1,u_2 \in L_{lin,kal}} \limsup_{N\to\infty} \frac{1}{N} \sum_{0\leq n<N} \mathbb{E}[qx^2[n] + r_1 u_1^2[n] + r_2 u_2^2[n]]}{\inf_{u_1,u_2} \limsup_{N\to\infty} \frac{1}{N} \sum_{0\leq n<N} \mathbb{E}[qx^2[n] + r_1 u_1^2[n] + r_2 u_2^2[n]]} \leq c.$$

Proof. See Appendix 10.2. This basically follows from the cases above. □

5.4.2 When $|a| = 1$

In this case, we can prove the following lemmas which parallel Lemma 5.3 and Corollary 5.4 from the case when $1 < |a| \leq 2.5$.

Lemma 5.4. *We use the definition of S_L shown in Lemma 5.3, i.e. the set of (k_1, k_2, k) such that $k_1, k_2, k \in \mathbb{N}$ and $1 \leq k_1 \leq k_2 \leq k$. We define $D_{L,2}(\widetilde{P_1}, \widetilde{P_2}, k_1, k_2, k)$ as follows:*

$$D_{L,2}(\widetilde{P_1}, \widetilde{P_2}, k_1, k_2, k) \geq (\sqrt{\frac{\Sigma + k_2 - k_1}{2^{2I'(P_1)}} + k - k_2} - \sqrt{(k-k_1)^2\widetilde{P_1}} - \sqrt{(k-k_2)^2\widetilde{P_2}})_+^2 + 1$$

where

$$\Sigma = \frac{k_1 - 1}{2^{2I}}$$
$$I = \frac{k_1 - 1}{2}\log(1 + \frac{k_1 - 1}{\sigma_{v1}^2}) + \frac{k_1 - 1}{2}\log(1 + \frac{k_1 - 1}{\sigma_{v2}^2})$$
$$I'(\widetilde{P_1}) = \frac{k_2 - k_1}{2}\log(1 + \frac{1}{\sigma_{v2}^2}(2\Sigma + 2(k_2 - k_1) + 2(k_2 - k_1 - 1)(k - k_1)\widetilde{P_1})).$$

Here, when $k_1 - 1 = 0$, $I = 0$ and when $k_2 - k_1 = 0$, $I'(\widetilde{P_1}) = 0$.

Let $|a| = 1$. Then, for all $q, r_1, r_2, \sigma_0, \sigma_{v1}, \sigma_{v2} \geq 0$, the minimum cost (5.1) of Problem J is lower bounded as follows:

$$\inf_{u_1,u_2} \limsup_{N\to\infty} \frac{1}{N} \sum_{0\leq n<N} q\mathbb{E}[x^2[n]] + r_1\mathbb{E}[u_1^2[n]] + r_2\mathbb{E}[u_2^2[n]]$$
$$\geq \sup_{(k_1,k_2,k)\in S_L} \min_{\widetilde{P_1},\widetilde{P_2}\geq 0} qD_{L,2}(\widetilde{P_1}, \widetilde{P_2}; k_1, k_2, k) + r_1\widetilde{P_1} + r_2\widetilde{P_2}.$$

Proof. See Appendix 10.3. This is similar to the proof of Lemma 5.3 except that the geometric sequences α_i and β_i in the geometric slicing are replaced by arithmetic sequences. □

Corollary 5.5. *Consider the decentralized LQG problem of Problem J. Let $|a| = 1$. Then, for all $q, r_1, r_2 > 0$, the minimum cost (5.1) of Problem J is lower bounded as follows:*

$$\inf_{u_1,u_2} \limsup_{N\to\infty} \frac{1}{N} \sum_{0\leq n<N} q\mathbb{E}[x^2[n]] + r_1\mathbb{E}[u_1^2[n]] + r_2\mathbb{E}[u_2^2[n]]$$
$$\geq \min_{\widetilde{P_1},\widetilde{P_2}\geq 0} qD_L(\widetilde{P_1}, \widetilde{P_2}) + r_1\widetilde{P_1} + r_2\widetilde{P_2}$$

where $D_L(\widetilde{P_1}, \widetilde{P_2})$ *satisfies the following conditions.*
(a) If $\sigma_{v2} \geq 16$ *and* $\widetilde{P_1} \leq \frac{1}{4\sigma_{v2}}$ *then* $D_L(\widetilde{P_1}, \widetilde{P_2}) \geq 0.09168\sigma_{v2} + 1$.
(b) If $\sigma_{v2} \geq 16$ *and* $\frac{1}{4\sigma_{v2}} \leq \widetilde{P_1} \leq \frac{1}{64}$ *then* $D_L(\widetilde{P_1}, \widetilde{P_2}) \geq \frac{0.02417}{\widetilde{P_1}} + 1$.
(c) If $\widetilde{P_1} \leq \frac{1}{50}$, $\widetilde{P_2} \leq \frac{1}{50}$ *then* $D_L(\widetilde{P_1}, \widetilde{P_2}) \geq \frac{0.003772}{\max(\widetilde{P_1}, \widetilde{P_2})} + 1$.
(d) For all $\widetilde{P_1}, \widetilde{P_2}$, $D_L(\widetilde{P_1}, \widetilde{P_2}) \geq \max(\frac{\sqrt{2}}{2}\sigma_{v1}, 1)$.

Proof. See Appendix 10.3. □

Like Section 5.4.1, we will intuitively argue why the power-distortion tradeoff can be characterized within a constant ratio by considering $D_L(\widetilde{P_1}, \widetilde{P_2})$ as if it is a lower bound on $D(P_1, P_2)$.

Notice that by (5.2), when $|a| = 1$ the Kalman filtering performance of the controllers are given as $\Sigma_1 = \frac{-1+\sqrt{1+4\sigma_v^2}}{2}$ and $\Sigma_2 = \frac{-1+\sqrt{1+4\sigma_v^2}}{2}$ respectively. Therefore, we can see $\Sigma_1 \approx \sigma_1$ and $\Sigma_2 \approx \sigma_1$ and so we can think of σ_1, σ_2 shown in Corollary 5.5 as if they are Σ_1, Σ_2.

As we discussed in Section 5.3.1, when $|a| = 1$ there is only one case for the power-distortion tradeoff. Thus, we will only divide the cases by P_1 and P_2.

• When $P_1 \leq \Theta(\frac{1}{\max(\Sigma_2,1)})$ and $P_2 \leq \Theta(\frac{1}{\max(\Sigma_2,1)})$. The controller with a larger power dominates the performance, and Figure 5.2 shows $D(P_1, P_2) = O(\frac{1}{\max(P_1,P_2)})$ is achievable. The statement (c) of Corollary 5.5 gives a matching lower bound.

• When $P_1 \leq \Theta(\frac{1}{\max(\Sigma_2,1)})$ and $\Theta(\frac{1}{\max(\Sigma_2,1)}) \leq P_2$. In this case, the second controller dominates the performance, but its performance saturates to the Kalman filtering performance. Figure 5.2 shows $D(P_1, P_2) = O(\max(\Sigma_2, 1))$. The statement (a) of Corollary 5.5 gives a matching lower bound.

• When $\Theta(\frac{1}{\max(\Sigma_2,1)}) \leq P_1 \leq \Theta(\frac{1}{\max(\Sigma_1,1)})$. In this case, the first controller dominates the performance, and Figure 5.2 shows $D(P_1, P_2) = O(\frac{1}{P_1})$. The statement (b) of Corollary 5.5 gives a matching lower bound.

• When $P_1 \geq \Theta(\frac{1}{\max(\Sigma_1,1)})$. In this case, the first controller dominates the performance, but its performance saturates to the Kalman filtering performance. Figure 5.2 shows $D(P_1, P_2) = O(\max(\Sigma_1, 1))$ is achievable. The statement (d) of Corollary 5.5 gives a matching lower bound.

Formally, the constant-ratio result for the average cost LQG problem can be written as follows.

Proposition 5.2. *Consider the decentralized LQG control of Problem J. There exists* $c \leq 540$ *such that for all* $|a| = 1$, q, r_1, r_2, σ_{v1} *and* σ_{v2},

$$\frac{\inf\limits_{u_1,u_2 \in L_{lin,kal}} \limsup\limits_{N\to\infty} \frac{1}{N} \sum\limits_{0\leq n<N} \mathbb{E}[qx^2[n] + r_1 u_1^2[n] + r_2 u_2^2[n]]}{\inf\limits_{u_1,u_2} \limsup\limits_{N\to\infty} \frac{1}{N} \sum\limits_{0\leq n<N} \mathbb{E}[qx^2[n] + r_1 u_1^2[n] + r_2 u_2^2[n]]} \leq c.$$

Proof. See Appendix 10.3. □

5.4.3 When $0.9 \leq |a| < 1$

Lemma 5.5. *We use the definition of S_L shown in Lemma 5.3, i.e. the set of (k_1, k_2, k) such that $k_1, k_2, k \in \mathbb{N}$ and $1 \leq k_1 \leq k_2 \leq k$. We define $D_{L,3}(\widetilde{P_1}, \widetilde{P_2}, k_1, k_2, k)$ as follows:*

$$D_{L,3}(\widetilde{P_1}, \widetilde{P_2}) := (\sqrt{\frac{\Sigma + a^{2(k-k_2+1)}\frac{1-a^{2(k_2-k_1)}}{1-a^2}}{2^{2I'(\widetilde{P_1})}} + a^2\frac{1-a^{2(k-k_2)}}{1-a^2}}$$
$$- \sqrt{(\frac{1-a^{k-k_1}}{1-a})^2\widetilde{P_1}} - \sqrt{(\frac{1-a^{k-k_2}}{1-a})^2\widetilde{P_2}})_+^2 + 1$$

where

$$\Sigma = \frac{a^{2(k-k_1+1)}\frac{1-a^{2(k_1-1)}}{1-a^2}}{2^{2I}}$$
$$I = \frac{1}{2}\log(1 + \frac{1}{\sigma_{v1}^2}\frac{1-a^{2(k_1-1)}}{1-a^2})^{k_1-1} + \frac{1}{2}\log(1 + \frac{1}{\sigma_{v2}^2}\frac{1-a^{2(k_1-1)}}{1-a^2})^{k_1-1}$$
$$I'(\widetilde{P_1}) = \frac{1}{2}\log(1 + \frac{1}{(k_2-k_1)\sigma_{v2}^2}(2a^{2(k_1-k)}\frac{1-a^{2(k_2-k_1)}}{1-a^2}\Sigma + 2(k_2-k_1)\frac{1-a^{2(k_2-1-k_1+1)}}{1-a^2}$$
$$+ 2a^{k_1-k}\frac{1-a^{k_2-k_1}}{1-a}\frac{(1-a^{k_2-1-k_1})(1-a^{k-k_1})}{(1-a)^2}\widetilde{P_1}))^{k_2-k_1}$$

Here, when $k_1 - 1 = 0$, $I = 0$ and when $k_2 - k_1 = 0$, $I'(\widetilde{P_1}) = 0$.

Let $0 \leq |a| < 1$. Then, for all $q, r_1, r_2, \sigma_0, \sigma_{v1}, \sigma_{v2} \geq 0$, the minimum cost (5.1) of Problem J is lower bounded as follows:

$$\inf_{u_1,u_2} \limsup_{N\to\infty} \frac{1}{N} \sum_{0\leq n<N} q\mathbb{E}[x^2[n]] + r_1\mathbb{E}[u_1^2[n]] + r_2\mathbb{E}[u_2^2[n]]$$
$$\geq \sup_{(k_1,k_2,k)\in S_L} \min_{\widetilde{P_1},\widetilde{P_2}\geq 0} qD_{L,3}(\widetilde{P_1}, \widetilde{P_2}; k_1, k_2, k) + r_1\widetilde{P_1} + r_2\widetilde{P_2}.$$

Proof. See Appendix 10.4 for the proof. The proof parallels to Lemma (5.3) except that the base for the geometric sequences α_i and β_i is a instead of a^{-1}. □

Corollary 5.6. *Consider the decentralized LQG problem of Problem J. Define*

$$\Sigma_1 := \frac{(a^2-1)\sigma_{v1}^2 - 1 + \sqrt{((a^2-1)\sigma_{v1}^2-1)^2 + 4a^2\sigma_{v1}^2}}{2a^2}$$
$$\Sigma_2 := \frac{(a^2-1)\sigma_{v2}^2 - 1 + \sqrt{((a^2-1)\sigma_{v2}^2-1)^2 + 4a^2\sigma_{v2}^2}}{2a^2}.$$

Let $0.9 \leq |a| < 1$. Then, for all $q, r_1, r_2 > 0$, the minimum cost (5.1) of Problem J is lower bounded as follows:

$$\inf_{u_1,u_2} \limsup_{N\to\infty} \frac{1}{N} \sum_{0\leq n<N} q\mathbb{E}[x^2[n]] + r_1\mathbb{E}[u_1^2[n]] + r_2\mathbb{E}[u_2^2[n]]$$
$$\geq \min_{\widetilde{P_1},\widetilde{P_2}\geq 0} qD_L(\widetilde{P_1}, \widetilde{P_2}) + r_1\widetilde{P_1} + r_2\widetilde{P_2}$$

where $D_L(\widetilde{P_1}, \widetilde{P_2})$ *satisfies the following conditions.*

Then, we have a lower bound $D_L(\widetilde{P_1}, \widetilde{P_2})$ *on* $D(P_1, P_2)$ *where* $D_L(\widetilde{P_1}, \widetilde{P_2})$ *satisfies the followings:*

(a) If $\Sigma_2 \geq 40$, $\widetilde{P_1} \leq \frac{1}{\Sigma_2}$ *then* $D_L(\widetilde{P_1}, \widetilde{P_2}) \geq 0.009131\Sigma_2 + 1$.

(b) If $\Sigma_2 \geq 40$, $\frac{1}{\Sigma_2} \leq \widetilde{P_1} \leq \frac{1}{40}$ *then* $D_L(\widetilde{P_1}, \widetilde{P_2}) \geq \frac{0.009131}{\widetilde{P_1}} + 1$.

(c) If $\frac{1-a^2}{20} \leq \max(\widetilde{P_1}, \widetilde{P_2}) \leq \frac{1}{40}$ *then* $D_L(P_1, P_2) \geq \frac{0.001201}{\max(\widetilde{P_1}, \widetilde{P_2})} + 1$.

(d) If $\max(\widetilde{P_1}, \widetilde{P_2}) \leq \frac{1-a^2}{20}$ *then* $D_L(\widetilde{P_1}, \widetilde{P_2}) \geq \frac{0.0869}{1-a^2} + 1$.

(e) For all $\widetilde{P_1}, \widetilde{P_2}$, $D_L(\widetilde{P_1}, \widetilde{P_2}) \geq \max(0.2636\Sigma_1, 1)$.

Proof. See Appendix 10.4 for the proof. □

Like in Section 5.4.1, we will intuitively argue why the power-distortion tradeoff can be characterized within a constant ratio by considering $D_L(\widetilde{P_1}, \widetilde{P_2})$ as if it is a lower bound on $D(P_1, P_2)$. The characterization of the power-distortion tradeoff is equivalent to the characterization of the average cost. Thus, we will intuitively argue how we can characterize the power-distortion tradeoff within a constant ratio. For this, we will first divide the cases by Σ_1, Σ_2, then we will further divide the cases by P_1, P_2.

When $\max(\Sigma_1, 1) \leq \max(\Sigma_2, 1) \leq \Theta(\frac{1}{1-a^2})$

We will further divide the cases based on $\widetilde{P_1}, \widetilde{P_2}$.

• When $P_1 \leq \Theta(1-a^2)$ and $P_2 \leq \Theta(1-a^2)$. From Figure 5.6a we can see that $D(P_1, P_2) = \frac{1}{1-a^2}$ is achievable without any control input. The statement (d) of Corollary 5.6 gives a lower bound tight up to a constant ratio.

• When $P_1 \leq \Theta(1-a^2)$ and $\Theta(1-a^2) \leq P_2 \leq \Theta(\frac{1}{\max(\Sigma_2,1)})$. The second controller dominates. Figure 5.6a shows $D(P_1, P_2) = O(\frac{1}{P_2})$ is achievable in this case. The statement (c) of Corollary 5.6 gives a lower bound tight up to a constant ratio.

• When $P_1 \leq \Theta(1-a^2)$ and $\Theta(\frac{1}{\max(\Sigma_2,1)}) \leq P_2$. In this case, the second controller's performance saturates to Kalman filtering, and Figure 5.6a shows $D(P_1, P_2) = O(\max(\Sigma_2, 1))$ is achievable. The statement (a) of Corollary 5.6 gives a lower bound tight up to a constant ratio.

• When $\Theta(1-a^2) \leq P_1 \leq \Theta(\frac{1}{\max(\Sigma_2,1)})$ and $P_2 \leq \Theta(1-a^2)$. The first controller dominates. Figure 5.6a shows $D(P_1, P_2) = O(\frac{1}{P_1})$ is achievable in this case. The statement (c) of Corollary 5.6 gives a lower bound tight up to a constant ratio.

• When $\Theta(1-a^2) \leq P_1 \leq \Theta(\frac{1}{\max(\Sigma_2,1)})$ and $\Theta(1-a^2) \leq P_2 \leq \Theta(\frac{1}{\max(\Sigma_2,1)})$. The controller with larger power dominates. Figure 5.6a shows $D(P_1, P_2) = O(\frac{1}{\max(P_1,P_2)})$ is achievable with the controller with lager power. The statement (c) of Corollary 5.6 gives a lower bound tight up to a constant ratio.

• When $\Theta(1-a^2) \leq P_1 \leq \Theta(\frac{1}{\max(\Sigma_2,1)})$ and $\Theta(\frac{1}{\max(\Sigma_2,1)}) \leq P_2$. The second controller dominates. Figure 5.6a shows $D(P_1, P_2) = O(\max(\Sigma_2, 1))$ is achievable. The statement (a) of Corollary 5.6 gives a lower bound tight up to a constant ratio.

• When $\Theta(\frac{1}{\max(\Sigma_2,1)}) \leq P_1 \leq \Theta(\frac{1}{\max(\Sigma_1,1)})$. The first controller dominates. Figure 5.6a shows $D(P_1,P_2) = O(\frac{1}{P_1})$ is achievable. The statement (b) of Corollary 5.6 gives a lower bound tight up to a constant ratio.

• When $\Theta(\frac{1}{\max(\Sigma_1,1)}) \leq P_1$. The first controller dominates. Figure 5.6a shows $D(P_1,P_2) = O(\max(\Sigma_1,1))$ is achievable. The statement (e) of Corollary 5.6 gives a lower bound tight up to a constant ratio.

When $\max(\Sigma_1,1) \leq \Theta(\frac{1}{1-a^2}) = \max(\Sigma_2,1)$

We will further divide the cases based on P_1, P_2.

• When $P_1 \leq \Theta(1-a^2)$. From Figure 5.6a we can see that $D(P_1,P_2) = \frac{1}{1-a^2}$ is achievable without any control input. The statement (b) of Corollary 5.6 gives a matching lower bound. More precisely, since $\Theta(\frac{1}{1-a^2}) = \max(\Sigma_2,1)$, for a large value of Σ_2 we can put $P_1 = \Theta(1-a^2)$ in the statement (b). Then, the bound reduces to $D(P_1,P_2) = \Omega(\frac{1}{1-a^2})$.

• When $\Theta(1-a^2) \leq P_1 \leq \Theta(\frac{1}{\max(\Sigma_1,1)})$. In this case, the first controller dominates, and Figure 5.6a shows $D(P_1,P_2) = O(\frac{1}{P_1})$ is achievable. The statement (b) of Corollary 5.6 gives a matching lower bound.

• When $\Theta(\frac{1}{\max(\Sigma_1,1)}) \leq P_1$. In this case, the first controller dominates, and Figure 5.6a shows $D(P_1,P_2) = O(\max(\Sigma_1,1))$ is achievable. The statement (e) of Corollary 5.6 gives a matching lower bound.

When $\Theta(\frac{1}{1-a^2}) = \max(\Sigma_1,1) \approx \max(\Sigma_2,1)$

In this case, the Kalman filtering noise Σ_1 and Σ_2 is already compatible with $\frac{1}{1-a^2}$, the state distortion attainable without any control inputs. Therefore, we cannot expect a significant control gain, and the optimal state distortion is $\Theta(\frac{1}{1-a^2})$. Since $\Theta(\frac{1}{1-a^2}) = \max(\Sigma_1,1)$, the statement (e) of Corollary 5.6 gives a matching lower bound.

Formally, the average LQG cost for $0.9 \leq |a| < 1$ can be characterized as follows.

Proposition 5.3. *Consider the decentralized LQG control of Problem J. There exists* $c \leq 1700$ *such that for all* $0.9 \geq |a| < 1$, q, r_1, r_2, σ_{v1} *and* σ_{v2},

$$\frac{\inf_{u_1,u_2 \in L_{lin,kal}} \limsup_{N\to\infty} \frac{1}{N} \sum_{0\leq n<N} \mathbb{E}[qx^2[n] + r_1 u_1^2[n] + r_2 u_2^2[n]]}{\inf_{u_1,u_2} \limsup_{N\to\infty} \frac{1}{N} \sum_{0\leq n<N} \mathbb{E}[qx^2[n] + r_1 u_1^2[n] + r_2 u_2^2[n]]} \leq c.$$

Proof. See Appendix 10.4 for the proof. □

5.4.4 When $|a| \leq 0.9$

Proposition 5.4. *Consider the decentralized LQG control of Problem J. There exists $c \leq 6$ such that for all $|a| < 0.9$, q, r_1, r_2, σ_{v1} and σ_{v2},*

$$\frac{\inf_{u_1,u_2 \in L_{lin,kal}} \limsup_{N\to\infty} \frac{1}{N} \sum_{0 \leq n < N} \mathbb{E}[qx^2[n] + r_1 u_1^2[n] + r_2 u_2^2[n]]}{\inf_{u_1,u_2} \limsup_{N\to\infty} \frac{1}{N} \sum_{0 \leq n < N} \mathbb{E}[qx^2[n] + r_1 u_1^2[n] + r_2 u_2^2[n]]} \leq c$$

Proof. By Lemma 4.14 of Chapter 4, it is enough to show that there exists $c \in \mathbb{R}$ such that $D_U(cP_1, cP_2) \leq c \cdot D_L(P_1, P_2)$.

Upper bound: Putting $k = 0$ in Lemma 5.1 gives

$$(D_U(P_1), P_1) \leq (\frac{1}{1-a^2}, 0) \leq (\frac{1}{1-0.9^2}, 0)$$

Lower bound: By Lemma 5.5,

$$D_L(P_1, P_2) \geq 1$$

Ratio: c is upper bounded by

$$c \leq \frac{1}{1-0.9^2} \leq 6$$

Therefore, the lemma is proved. □

5.4.5 Proof of Theorem 5.1

Now, by combining the results of Proposition 5.1, 5.2, 5.3 we can prove the main theorem of the chapter.

Proof of Theorem 5.4.5. The proof immediately follows from Proposition 5.1, 5.2, 5.3. □

Chapter 6

Conclusion

In this thesis, we studied modern control system problems through an informational lens. Even though we considered explicit control systems with control objectives, what we really found in the solutions was the nature of information flow. We applied information-theoretic ideas to understand the problems, often found relations between control and communication theory, and even discovered hints of a unified theory. Table 6.1 shows a list of ideas, results and techniques shown in this thesis, and Figure 6.1 summaries the relationship and parallelism between the two theories.

One of the earliest paper which showed an explicit relationship between communication and control is due to Schalkwijk and Kailath. In [89], they found a surprising relationship between point-to-point feedback communication systems and centralized feedback control systems. Recently, in [37] we found another interesting relationship between dirty-paper coding (in information theory) [20] and Witsenhausen's counterexample (in control theory) [108]. We proposed a control strategy based on the known solution for dirty-paper coding, and applied large deviation ideas [24] to prove its approximate optimality.

In this thesis, we found much more extensive relationships between the two theories. In Chapter 2, we considered intermittent Kalman filtering, which previously had been considered only from a control theoretic point of view. However, we found that the essence of the problem is in fact communication and information flows. The plant in intermittent Kalman filtering can be thought of as the source of information flows, and the observability gramian generated from the successfully received observations can be thought of as the channel. Furthermore, we showed that the different subspaces of the plant do not interact with each other as long as they belong to different eigenvalue cycles. To justify this, we adapt successive decoding ideas [21] from information theory. Then, the amount of source information and the channel capacity were measured by rank, i.e. the dimension of subspaces belong to the same eigenvalue cycle is the amount of source information, and the rank of the observability gramian generated from the successfully received observations is the channel

	• Intermittent Kalman Filtering (Chapter 2)
Result	(1) Characterization of critical erasure probability (Theorem 2.7 of Page 39)
Idea	- Eigenvalue Cycle to capture periodicity of system (Section 2.5.1 of Page 30)
SP Technique	- Polyphase decomposition idea to reduce periodic systems to aperiodic systems (Claim 7.7 of Page 328)
IT Technique	- Successive decoding idea to justify one state by one state decoding (Section 2.5.2 of Page 33)
IT Technique	- Large deviation idea to analyze the p.m.f. tails (Appendix 7.1 of Page 260)
Ergodic Theory	- Weyl's criterion to approximate deterministic sequences by random variables (Appendix 7.6 of Page 300)
Strategy Result	(2) Nonuniform sampling improves performance (Theorem 2.8)
IT Technique	- Random jittering idea for analysis (Section 2.7 of Page 43)
Technique	- Application of Dini's theorem to justify uniform convergence (Appendix 7.3 of Page 272)
	• Network coding meets Decentralized control (Chapter 3)
Result	(1) Algebraic mincut-maxflow theorem (Theorem 3.2 of Page 85)
Idea	- Network Linearization idea to simplify network topology to single-hop relay networks (Section 3.2.2 of Page 86)
Interpretation Result	(2) Externalization of Implicit Communication (Section 3.5 of Page 110)
Idea	- Jordan form transition to reveal source and destination of information flow (Section 3.5.2 of Page 113)
Idea & Technique	- Interpretation of transfer function as network topology (Theorem 3.8 of Page 121)
New Problem	(3) Control over LTI networks (Section 3.6 of Page 3.6)
Result	- Relationship between network capacity and stabilizability of system (Theorem 3.9 of Page 124)
Result	- Use of multicast network coding schemes to increase reliability control systems (Theorem 3.10 of Page 130)
Result	- Use of broadcast network coding schemes to reduce interference between control systems (Theorem 3.12 of Page 135)
	• Scalar LQG problem with two controllers - Fast Dynamics (Chapter 4)
Strategy Result	(1) Achievable Cost (Lemma 4.7 of Page 174)
IT Technique	- Linear determinist model interpretation (Section 4.3 of Page 148)
Idea	- Nonlinear s-stage signaling strategy (Definition 4.2 of Page 145)
IT Technique	- Approximate lattice theory to analyze the performance (Section 4.5.2 of Page 170)
Result	(2) Lower bound on Cost (Lemma 4.12 of Page 198)
Technique	- Geometric slicing to reduce infinite-horizon problems to finite-horizon problems (Lemma 4.8 of Page 180)
Technique	- Three stage division of finite-horizon problems (Section 4.6.2 of Page 185)
Technique	- Bounding the first interval as information limited interval (Lemma 4.9 of Page 187)
Result & Technique	- Interpretation of the second interval as MIMO Witsenhausen's counterexample (Section 4.6.3 of Page 190)
Technique	- Use of large deviation ideas to bound rare events (Proof of Lemma 4.12 of Page 198)
Technique	- Bounding the third interval as power limited interval (Lemma 4.10 of Page 188)
Result	(3) Constant Ratio Optimality Result (Theorem 4.1 of Page 146)
Technique	- Relation between average cost problem and power-distortion tradeoff (Lemma 4.14 of Page 212)
	• Scalar LQG problem with two controllers - Slow Dynamics (Chapter 5)
Result	(1) Achievable Cost (Lemma 5.3 of Page 227)
Interpretation & Result	- Approximation of centralized control cost (Corollary 5.1, 5.2, 5.3 of Page 230, 233, 235)
Result	(2) Lower bound on Cost (Section 5.4 of Page 236)
Technique	- Geometric slicing with different sequences (Lemma 5.2 of Page 238)

Table 6.1: Highlighted Ideas, Results and Techniques developed in this thesis

capacity. This insight parallels the fact that in MIMO AWGN(additive white Gaussian noise) communication channels the ranks of channel matrices are known as the d.o.f.(degree-of-freedom) capacity of the channel [99].

Therefore, the intermittent Kalman filtering performance is deeply related with the rank of the observability gramian generated from the randomly received observations. We saw that the probability that such randomly generated observability gramians have too small a rank to convey enough information about the plant dominates the performance of intermittent Kalman filtering. We adapted large-deviation ideas [24] from information theory to analyze such a probability. Furthermore, nonuniform sampling can be used as a simple way to increase the rank of the observability gramian, so it can dramatically improve the intermittent Kalman filtering performance.

In Chapter 3, we took a unified view of distributed linear control systems and linear network coding. By restricting the system, controllers, transmitter, relays and receiver designs to be linear time-invariant, we considered both systems as linear time-invariant systems. Based on this interpretation, we developed an algorithm which extracts implicit information flows that must happen when the controllers stabilize the plant. More precisely, we modeled the implicit information flows by relay networks. The source and destination of the relay network are the states(subspaces) of the plant corresponding to the same eigenvalue, the relays are the controllers, and the remaining states of the plant correspond to the channels. Like in the intermittent Kalman filtering interpretation, information is measured by a rank. The dimension of subspaces corresponds to the minimum amount of information that must flow to stabilize them. Therefore, a subspaces can be stabilized if and only if the mincut of the relay network is larger or equal to this minimum required amount of information. Here, the mincut of the relay network is also measured by a rank, precisely the rank of the channel matrix for the cut. Thus, we could understand the stabilizability condition for distributed control systems through the lens of a mincut-maxflow theorem for relay communication networks. We also saw that this insight can lead to new designs for distributed control with explicit LTI communication networks.

Furthermore, the connection between distributed control and network coding could lead to new results for network coding. In Chapter 3.2.2, by applying state-space representation ideas to network coding, we found an algorithm that converts arbitrary topology communication networks to equivalent single-hop relay networks, which we called network linearization. This standardization of network topology turned out to be extremely useful when the complexity of network topology is the crux of the problem. We asked the question whether the mincuts of LTI networks are achievable by static LTI relay schemes. We first showed the answer is yes for standardized single-hop relay networks. Then, we generalized the result to arbitrary topology networks by network linearization.

Finally, in Chapter 4 and 5, we considered the optimal LQG control problem with a scalar plant and two controllers, and leveraged the understanding of control information flows to approximately optimal controller design. One of the key ideas in finding an approximate optimal strategy was an appropriate division of cases. Just as wireless communication theory [99] divides cases according to SNR (signal-to-noise ratio), we divided based on the eigenvalue of the system. When the eigenvalue of the system is large, we called it the fast-dynamics case. When the eigenvalue of the system is small, we called it the slow-dynamics case.

The main insight to understand fast-dynamics cases was a linear view of nonlinearity. We saw that in the fast-dynamics case, nonlinear controllers can infinitely outperform linear controllers. To understand nonlinear controllers and resulting nonlinear system's behavior, we considered each bit-level of the state as different linear spaces. In the resulting linear deterministic model [6], information still can be measured by a rank. In the proposed approximately optimal nonlinear strategy, the first controller "communicates" to the second controller by reducing the rank of the binary representation of the state. In other words, by reducing the rank of the binary representation, the first controller reduces the amount of information in the state so that the second controller can have better estimates about the state. Since we are focusing on the rank of linear spaces, this control strategy parallels with high-SNR wireless communication schemes which exploit d.o.f. gain [99].

For slow-dynamics, we saw that the opposite is true. The SNR of implicit communication between two controllers is bounded by the eigenvalue. Therefore, there is no huge incentive for implicit communication, and single controller linear strategy (Kalman filtering) turns out to be approximately optimal. The Kalman filtering gain can be thought of as a kind of power gain which turned out to be crucial for low-SNR case [99] in wireless communication.

Furthermore, to prove approximate optimality of the proposed strategy, we used information theory and found new fundamental limits on control performance. A key tool was the geometric slicing idea, which gives different ways of cutting infinite-horizon problems into finite-horizon ones, and parallels with cutset bounds in information theory.

Control and communication theory have been developed separately for decades. However, as we saw in this thesis, there exist extensive parallelism and relationship between two theories. Furthermore, the emerging modern systems have both control and communication systems as subsystems. A mathematical theory for modern cyber-physical systems should include control and communication theories. Figure 6.1 summaries the relationships between control and communication theories, and also indicates the directions to build a unified theory for modern systems. Lots of ideas and problems in both theories still remain unconnected, and these connections have to be made to build modern systems.

As we control power grids or transportation systems over communication networks, the security issue becomes a crucial component. The security concept of communication theory has to be connected to the safety concept of control theory. As modern systems scale, we also have to understand overall behavior of whole system as the number of subsystems grows. Since control systems keep evolving over time, the delay issue of communication is becoming more crucial and has to be theoretically understood.

All of these theoretic understanding and insights into modern systems have to based on the understanding of information flows for control. This thesis shows the possibility of theoretic study

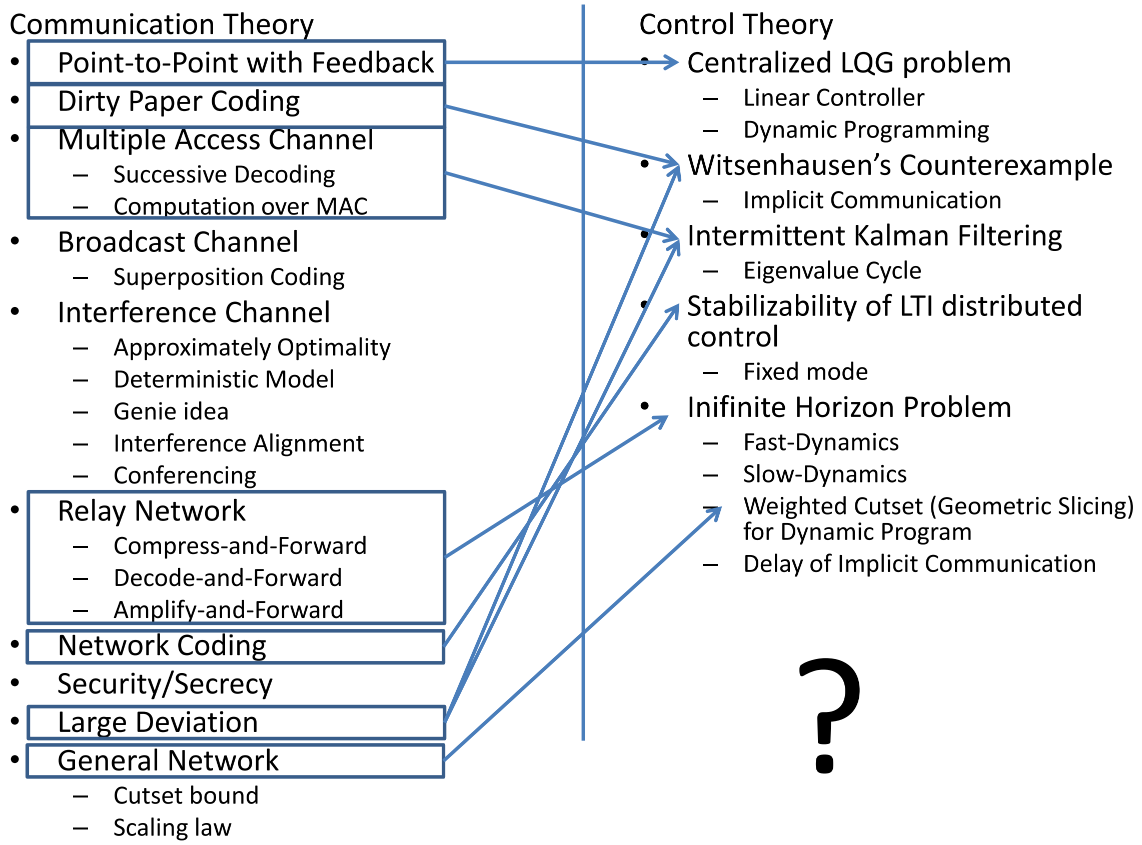


Figure 6.1: Parallelism and Relationship between Communication and Control Theory

of control information flows. Even thought control information flows have their own unique features, they still bears lots of semblance to wireless information flows. This suggests that we can exploit the current understanding of wireless information flows to study control information flows. At the end of study, we may be able to reach a unified theory for control and communication from a lens of information.

Chapter 7

Appendix for Chapter 2

7.1 Lemmas for Tails of Probability Mass Functions

In this section, we will prove some properties of the tails of probability mass functions (p.m.f.). By the tail, we mean how fast the probability decreases geometrically as we consider rarer and rarer events.

First, we define the essential supremum, $\operatorname{ess\,sup}$.

Definition 7.1. *For a given random variable X, $\operatorname{ess\,sup} X$ is given as follows.*

$$\operatorname{ess\,sup} X = \inf\{x \in \mathbb{R} : \mathbb{P}(X > x) = 0\}.$$

The following lemma shows that even if we increase a random variable sub-linearly, its p.m.f. tail remains the same.

Lemma 7.1. *Consider σ-field $\mathcal{F}$ and a nonnegative discrete random variable k whose probability mass function satisfies*

$$\exp(\limsup_{n\to\infty} \operatorname{ess\,sup} \frac{1}{n} \log \mathbb{P}\{k = n|\mathcal{F}\}) \le p$$

Then, given a function $f(x)$ such that $f(x) \le a(\log(x+1)+1)$ for some $a \in \mathbb{R}^+$, the probability mass function of a random variable $k + f(k)$ satisfies the following:

$$\exp(\limsup_{n\to\infty} \operatorname{ess\,sup} \frac{1}{n} \log \mathbb{P}\{k + f(k) = n|\mathcal{F}\}) \le p.$$

Proof. Since $\operatorname{ess\,sup} \mathbb{P}\{k = n|\mathcal{F}\}$ is bounded by 1, for all $\delta > 0$ such that $p + \delta < 1$ we can find a positive c such that $\operatorname{ess\,sup} \mathbb{P}\{k = n|\mathcal{F}\} \le c\,(p+\delta)^n\,(1-(p+\delta))$. Moreover, since $f(x) \lesssim \log(x+1)+1$, for all $\delta' > 0$ we can find a positive c' such that $f(x) \le \delta' x + c'$ for all $x \in \mathbb{R}^+$. Then, we have

$$\begin{aligned}\operatorname{ess\,sup} \mathbb{P}\{k+f(k)=n|\mathcal{F}\} &\le \operatorname{ess\,sup} \mathbb{P}\{k+f(k)\ge n|\mathcal{F}\} \le \operatorname{ess\,sup} \mathbb{P}\{k+\delta' k+c' \ge n|\mathcal{F}\}\\ &\le \operatorname{ess\,sup} \mathbb{P}\{k \ge \lfloor \frac{n-c'}{1+\delta'} \rfloor |\mathcal{F}\} \le \sum_{i=\lfloor \frac{n-c'}{1+\delta'} \rfloor}^{\infty} \operatorname{ess\,sup} \mathbb{P}\{k=i|\mathcal{F}\}\\ &\le \sum_{i=\lfloor \frac{n-c'}{1+\delta'} \rfloor}^{\infty} c(p+\delta)^i(1-(p+\delta))\\ &= c(1-(p+\delta))\frac{(p+\delta)^{\lfloor \frac{n-c'}{1+\delta'} \rfloor}}{1-(p+\delta)} = c(p+\delta)^{\lfloor \frac{n-c'}{1+\delta'} \rfloor}\\ &\le c(p+\delta)^{\frac{n-c'}{1+\delta'}-1} = c(p+\delta)^{-\frac{c'}{1+\delta'}-1}(p+\delta)^{\frac{n}{1+\delta'}}.\end{aligned}$$

Therefore,

$$\exp\left(\limsup_{n\to\infty} \operatorname{ess\,sup} \frac{1}{n} \log \mathbb{P}\{k+f(k)=n|\mathcal{F}\}\right) \le (p+\delta)^{\frac{1}{1+\delta'}}.$$

Since we can choose δ and δ' arbitrarily close to 0,

$$\exp\left(\limsup_{n\to\infty} \operatorname{ess\,sup} \frac{1}{n} \log \mathbb{P}\{k+f(k)=n|\mathcal{F}\}\right) \le p,$$

which finishes the proof. □

The following (well-known) lemma tells us that if we add independent random variables, the p.m.f. tail of the sum is equal to the heaviest one.

Lemma 7.2. *Consider an increasing σ-field sequence $\mathcal{F}_0, \mathcal{F}_1, \cdots, \mathcal{F}_{n-1}$ and a sequence of discrete random variables $k_1, k_2, \cdots, k_n$ satisfying two properties:*
(i) $k_i \in \mathcal{F}_i$ for $i \in \{1, \cdots, n-1\}$
(ii) $\exp(\limsup_{k\to\infty} \operatorname{ess\,sup} \frac{1}{k} \log \mathbb{P}(k_i = k|\mathcal{F}_{i-1})) \le p_i$.
Let $S = \sum_{i=1}^n k_i$. Then, $\exp(\limsup_{s\to\infty} \operatorname{ess\,sup} \frac{1}{s} \log \mathbb{P}(S = s|\mathcal{F}_0)) \le \max_{1\le i\le n}\{p_i\}$.

Proof. Given $\delta > 0$, let k'_i be independent geometric random variables with probability $1-(p_i+\delta)$. Denote $S' := \sum_{i=1}^n k'_i$. The moment generating function of S' is

$$\mathbb{E}[Z^{-S'}] = \prod_{i=1}^n \frac{(1-(p_i+\delta))}{1-(p_i+\delta)\, Z^{-1}}.$$

By [75], the last term can be expanded into a sum of rational functions whose denominators are $1-(p_i+\delta)Z^{-1}$. Therefore, by using an inverse Z-transform shown in [75], we can prove that

$$\exp(\limsup_{s\to\infty} \frac{1}{s} \log \mathbb{P}(S' = s)) \le \max_{1\le i\le n}\{p_i+\delta\}.$$

On the other hand, since $\operatorname{ess\,sup} \mathbb{P}(k_i = k|\mathcal{F}_{i-1})$ is bounded by 1, for all $\delta > 0$ we can find positive c_i such that

$$\operatorname{ess\,sup} \mathbb{P}(k_i = k|\mathcal{F}_{i-1}) \le c_i\,(p_1+\delta)^k\,(1-(p_1+\delta)) = c_i \mathbb{P}(k'_i = k)$$

for all $k \in \mathbb{Z}^+$. Then

$$\begin{aligned}
&\operatorname{ess\,sup} \mathbb{P}(S = s|\mathcal{F}_0) \\
&= \operatorname{ess\,sup} \sum_{s=s_1+\cdots+s_n} \mathbb{P}(k_1 = s_1|\mathcal{F}_0)\mathbb{P}(k_2 = s_2|\mathcal{F}_0, k_1 = s_1)\cdots\mathbb{P}(k_n = s_n|\mathcal{F}_0, k_1 = s_1, \cdots, k_{n-1} = s_{n-1}) \\
&\le \sum_{s=s_1+\cdots+s_n} \operatorname{ess\,sup} \mathbb{P}(k_1 = s_1|\mathcal{F}_0) \operatorname{ess\,sup} \mathbb{P}(k_2 = s_2|\mathcal{F}_1)\cdots \operatorname{ess\,sup} \mathbb{P}(k_n = s_n|\mathcal{F}_{n-1}) \\
&\le \prod_{1\le i\le n} c_i \cdot \sum_{s=s_1+\cdots+s_n} \mathbb{P}(k'_1 = s_1)\mathbb{P}(k'_2 = s_2)\cdots\mathbb{P}(k'_n = s_n) \\
&\le \prod_{1\le i\le n} c_i \cdot \mathbb{P}(S' = s).
\end{aligned}$$

Thus, $\exp(\limsup_{s\to\infty} \operatorname{ess\,sup} \frac{1}{s} \log \mathbb{P}(S = s|\mathcal{F}_0)) \le \max_{1\le i\le n}\{p_i + \delta\}$.

Since this holds for all $\delta > 0$, $\exp(\limsup_{s\to\infty} \operatorname{ess\,sup} \frac{1}{s} \log \mathbb{P}(S = s|\mathcal{F}_0)) \le \max_{1\le i\le n}\{p_i\}$. □

The next lemma tells us how the large deviation principle [24] can be applied to stopping times, i.e. it formally states the "test channel" and the "distance idea" shown in the power property of Section 2.5.1.

Lemma 7.3. *For given n, consider discrete random variables $k_1, k_2, \cdots, k_n$ and σ-algebra $\mathcal{F}$. The probability mass functions of $k_1, k_2 \cdots, k_n$ satisfy*

$$\exp(\limsup_{k\to\infty} \operatorname{ess\,sup} \frac{1}{k} \log \mathbb{P}\{k_i = k|\mathcal{F}\}) \le p_i$$

and $k_1, k_2, \cdots, k_n$ are conditionally independent given $\mathcal{F}$.
For given sets $T_1, T_2, \cdots, T_m \subseteq \{1, 2, \cdots, n\}$, define stopping times $M_1, \cdots, M_m$ as

$$M_i := \max_{t\in T_i} k_t$$

and a stopping time S as

$$S := \min_{1\le i\le m} M_i.$$

Then,

$$\exp\left(\limsup_{k\to\infty} \operatorname{ess\,sup} \frac{1}{k} \log \mathbb{P}\{S = k|\mathcal{F}\}\right) \le \max_{T=\{t_1,t_2,\cdots,t_{|T|}\}\subseteq\{1,2,\cdots,n\} \text{ s.t. } T\cap T_i\neq\emptyset \text{ for all } i} p_{t_1}p_{t_2}\cdots p_{t_{|T|}}.$$

Proof. Since $\operatorname{ess\,sup} \mathbb{P}\{k_i = k|\mathcal{F}\}$ is bounded by 1, for all $\delta > 0$ we can find $c > 1$ such that

$$\operatorname{ess\,sup} \mathbb{P}\{k_i = k|\mathcal{F}\} \le c(p_i + \delta)^k \left(1 - (p_i + \delta)\right).$$

Thus, we have

$$\operatorname{ess\,sup} \mathbb{P}\{k_i \ge k|\mathcal{F}\} \le c(p_i + \delta)^k.$$

Therefore,

$$
\begin{aligned}
\operatorname{ess\,sup} \mathbb{P}\{S = k|\mathcal{F}\} &\leq \operatorname{ess\,sup} \mathbb{P}\{S \geq k|\mathcal{F}\} \\
&= \operatorname{ess\,sup} \mathbb{P}\{M_1 \geq k, \cdots, M_m \geq k|\mathcal{F}\} \\
&= \operatorname{ess\,sup} \mathbb{P}\{\text{There exists } T = \{t_1, t_2, \cdots, t_{|T|}\} \subseteq \{1, \cdots, n\} \text{ s.t.} \\
&\quad T \cap T_i \neq \emptyset \text{ for all } i \text{ and } k_{t_1} \geq k, \cdots, k_{t_{|T|}} \geq k|\mathcal{F}\} \\
&\leq \sum_{\substack{T = \{t_1, t_2, \cdots, t_{|T|}\} \subseteq \{1, \cdots, n\} \\ \text{s.t. } T \cap T_i \neq \emptyset \text{ for all } i}} \operatorname{ess\,sup} \mathbb{P}\{k_{t_1} \geq k, k_{t_2} \geq k, \cdots, k_{t_{|T|}} \geq k|\mathcal{F}\} \\
&\leq |\{T = \{t_1, t_2, \cdots, t_{|T|}\} \subseteq \{1, \cdots, n\} \text{ s.t. } T \cap T_i \neq \emptyset \text{ for all } i\}| \qquad (7.1) \\
&\quad \cdot \max_{\substack{T = \{t_1, t_2, \cdots, t_{|T|}\} \subseteq \{1, \cdots, n\} \\ \text{s.t. } T \cap T_i \neq \emptyset \text{ for all } i}} \operatorname{ess\,sup} \mathbb{P}\{k_{t_1} \geq k|\mathcal{F}\} \cdots \operatorname{ess\,sup} \mathbb{P}\{k_{t_{|T|}} \geq k|\mathcal{F}\} \\
&\leq c^n |\{T = \{t_1, t_2, \cdots, t_{|T|}\} \subseteq \{1, \cdots, n\} \text{ s.t. } T \cap T_i \neq \emptyset \text{ for all } i\}| \\
&\quad \cdot \max_{\substack{T = \{t_1, t_2, \cdots, t_{|T|}\} \subseteq \{1, \cdots, n\} \\ \text{s.t. } T \cap T_i \neq \emptyset \text{ for all } i}} (p_{t_1} + \delta)^{k-1} (p_{t_2} + \delta)^{k-1} \cdots (p_{t_{|T|}} + \delta)^{k-1}.
\end{aligned}
$$

(7.1) follows from union bound. Since the above inequality holds for all $\delta > 0$,

$$\exp\left(\limsup_{k \to \infty} \operatorname{ess\,sup} \frac{1}{k} \log \mathbb{P}\{S = k|\mathcal{F}\}\right) \leq \max_{T = \{t_1, t_2, \cdots, t_{|T|}\} \subseteq \{1, 2, \cdots, n\} \text{ s.t. } T \cap T_i \neq \emptyset \text{ for all } i} p_{t_1} p_{t_2} \cdots p_{t_{|T|}}.$$

□

7.2 Lemmas about the Observability Gramian of Continuous-Time Systems

In linear system theory [17], the observability Gramian plays a crucial role in estimating states from observations. Therefore, we also study the behavior of the observability Gramian, especially the norm of the inverse of the observability Gramian.

First, we start with a corollary of the classic rearrangement inequality [43].

Lemma 7.4 (Rearrangement Inequality). *For $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 0$, $0 \leq k_1 \leq k_2 \leq \cdots \leq k_m$, and any permutation map σ, the following inequality is true:*

$$e^{-\lambda_{\sigma(1)} k_1} e^{-\lambda_{\sigma(2)} k_2} \cdots e^{-\lambda_{\sigma(m)} k_m} \leq e^{-\lambda_1 k_1} e^{-\lambda_2 k_2} \cdots e^{-\lambda_m k_m}.$$

Moreover, the ratio of these two can also be upper bounded as

$$\frac{e^{-\lambda_{\sigma(1)} k_1} e^{-\lambda_{\sigma(2)} k_2} \cdots e^{-\lambda_{\sigma(m)} k_m}}{e^{-\lambda_1 k_1} e^{-\lambda_2 k_2} \cdots e^{-\lambda_m k_m}} \leq e^{-(\lambda_{\sigma(m)} - \lambda_m)(k_m - k_{\sigma^{-1}(m)})}.$$

Proof. The first inequality directly follows from the classic rearrangement inequality. The second inequality is proved as follows: When $\sigma^{-1}(m) = m$, the inequality is trivial. When $\sigma^{-1}(m) \neq m$, we have

$$\begin{aligned}
&e^{-\lambda_{\sigma(1)}k_1} e^{-\lambda_{\sigma(2)}k_2} \cdots e^{-\lambda_m k_{\sigma^{-1}(m)}} \cdots e^{-\lambda_{\sigma(m-1)}k_{m-1}} e^{-\lambda_{\sigma(m)}k_m} \\
&= \underbrace{\left(e^{-\lambda_{\sigma(1)}k_1} e^{-\lambda_{\sigma(2)}k_2} \cdots e^{-\lambda_m k_{\sigma^{-1}(m)}} \cdots e^{-\lambda_{\sigma(m-1)}k_{m-1}}\right)}_{(a)} \cdot e^{-\lambda_{\sigma(m)}k_m} \\
&= \underbrace{\left(e^{-\lambda_{\sigma(1)}k_1} e^{-\lambda_{\sigma(2)}k_2} \cdots e^{-\lambda_{\sigma(m)} k_{\sigma^{-1}(m)}} \cdots e^{-\lambda_{\sigma(m-1)}k_{m-1}}\right)}_{(b)} \cdot \left(\frac{e^{-\lambda_m k_{\sigma^{-1}(m)}}}{e^{-\lambda_{\sigma(m)} k_{\sigma^{-1}(m)}}}\right) \cdot e^{-\lambda_{\sigma(m)}k_m}. \qquad (7.2)
\end{aligned}$$

We can notice that the exponent of (a) has $\{\lambda_1, \lambda_2, \cdots, \lambda_m\} \setminus \{\lambda_{\sigma(m)}\}$ and $\{k_1, k_2, \cdots, k_m\} \setminus \{k_m\}$ terms in it, and the exponent of (b) has

$$\begin{aligned}
&\left(\{\lambda_1, \lambda_2, \cdots, \lambda_m\} \setminus \{\lambda_{\sigma(m)}\}\right) \cup \{\lambda_{\sigma(m)}\} \setminus \{\lambda_m\} \\
&= \{\lambda_1, \lambda_2, \cdots, \lambda_m\} \setminus \{\lambda_m\}
\end{aligned}$$

and $\{k_1, k_2, \cdots, k_m\} \setminus \{k_m\}$ terms in it. Thus, by the first inequality of the lemma,

$$(b) \leq e^{-\lambda_1 k_1} \cdots e^{-\lambda_{m-1} k_{m-1}}.$$

Together with (7.2), we have

$$\begin{aligned}
&\frac{e^{-\lambda_{\sigma(1)}k_1} e^{-\lambda_{\sigma(2)}k_2} \cdots e^{-\lambda_{\sigma(m)}k_m}}{e^{-\lambda_1 k_1} e^{-\lambda_2 k_2} \cdots e^{-\lambda_m k_m}} \\
&\leq \frac{\left(e^{-\lambda_1 k_1} \cdots e^{-\lambda_{m-1} k_{m-1}}\right) \cdot \left(\frac{e^{-\lambda_m k_{\sigma^{-1}(m)}}}{e^{-\lambda_{\sigma(m)} k_{\sigma^{-1}(m)}}}\right) \cdot e^{-\lambda_{\sigma(m)}k_m}}{e^{-\lambda_1 k_1} e^{-\lambda_2 k_2} \cdots e^{-\lambda_m k_m}} \\
&= \frac{1}{e^{-\lambda_m k_m}} \cdot \left(\frac{e^{-\lambda_m k_{\sigma^{-1}(m)}}}{e^{-\lambda_{\sigma(m)} k_{\sigma^{-1}(m)}}}\right) \cdot e^{-\lambda_{\sigma(m)}k_m} = e^{(\lambda_m - \lambda_{\sigma(m)})(k_m - k_{\sigma^{-1}(m)})}
\end{aligned}$$

which finishes the proof. ☐

Even though Theorem 2.8 is written for a general matrix $\mathbf{C}$, we will first start from the simpler case of a row vector $\mathbf{C}$. In fact, for the proof of the general case, we will reduce the system with a matrix $\mathbf{C}$ to a system with a row vector $\mathbf{C}$.

First, we introduce the definitions corresponding to (2.28), (2.29) for a row vector $\mathbf{C}$. Let

$\mathbf{A_c}$ be a $m \times m$ Jordan form matrix, and $\mathbf{C}$ be a $1 \times m$ row vector $\mathbf{C}$ which are written as follows:

$$\mathbf{A_c} = diag\{\mathbf{A_{1,1}}, \mathbf{A_{1,2}}, \cdots, \mathbf{A_{1,\nu_1}}, \cdots, \mathbf{A_{\mu,1}}, \cdots, \mathbf{A_{\mu,\nu_\mu}}\} \tag{7.3}$$

$$\mathbf{C} = \begin{bmatrix} \mathbf{C_{1,1}} & \mathbf{C_{1,2}} & \cdots & \mathbf{C_{1,\nu_1}} & \cdots & \mathbf{C_{\mu,1}} & \cdots & \mathbf{C_{\mu,\nu_\mu}} \end{bmatrix} \tag{7.4}$$

where $\mathbf{A_{i,j}}$ is a Jordan block with eigenvalue $\lambda_{i,j} + \sqrt{-1}\omega_{i,j}$ and size $m_{i,j}$

$m_{i,1} \leq m_{i,2} \leq \cdots \leq m_{i,\nu_i}$ for all $i = 1, \cdots, \mu$

$m_i = \sum_{1 \leq j \leq \nu_i} m_{i,j}$ for all $i = 1, \cdots, \mu$

$\lambda_{i,1} = \lambda_{i,2} = \cdots = \lambda_{i,\nu_i}$ for all $i = 1, \cdots, \mu$

$\lambda_{1,1} > \lambda_{2,1} > \cdots > \lambda_{\mu,1} \geq 0$

$\omega_{i,1}, \cdots, \omega_{i,\nu_i}$ are pairwise distinct

$\mathbf{C_{i,j}}$ is a $1 \times m_{i,j}$ complex matrix and its first element is non-zero

$\lambda_i + \sqrt{-1}\omega_i$ is (i,i) element of $\mathbf{A_c}$.

Here, we can notice that the real parts of the eigenvalues of $\mathbf{A_{i,1}}, \cdots, \mathbf{A_{i,\nu_i}}$ are the same, but the eigenvalues of all Jordan blocks $\mathbf{A_{i,j}}$ are distinct. Therefore, by Theorem 2.6, the condition that the first elements of $\mathbf{C_{i,j}}$ are non-zero corresponds to the observability of $(\mathbf{A_c}, \mathbf{C})$.

The following lemma upper bounds the determinant of the observability Gramain of the sampled continuous system.

Lemma 7.5. *Let $\mathbf{A_c}$ and $\mathbf{C}$ be given as* (7.3) *and* (7.4)*. For $0 \leq k_1 \leq k_2 \leq \cdots \leq k_m$, there exists $a > 0$, $p \in \mathbb{Z}^+$ such that*

$$\left| \det\left(\begin{bmatrix} \mathbf{C}e^{-k_1\mathbf{A_c}} \\ \mathbf{C}e^{-k_2\mathbf{A_c}} \\ \vdots \\ \mathbf{C}e^{-k_m\mathbf{A_c}} \end{bmatrix} \right) \right| \leq a(k_m^p + 1) \prod_{1 \leq i \leq m} e^{-k_i \lambda_i}$$

where λ_i is the real part of (i,i) component of $\mathbf{A_c}$.

Proof. First consider a diagonal matrix, i.e. $\mathbf{A_c} = \begin{bmatrix} \lambda_1 + j\omega_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 + j\omega_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_m + j\omega_m \end{bmatrix}$. Then,

$$\begin{aligned}
&\left| \det \left(\begin{bmatrix} \mathbf{C}e^{-k_1 \mathbf{A_c}} \\ \mathbf{C}e^{-k_2 \mathbf{A_c}} \\ \vdots \\ \mathbf{C}e^{-k_m \mathbf{A_c}} \end{bmatrix} \right) \right| \\
&= \left| \sum_{\sigma \in S_m} sgn(\sigma) \prod_{i=1}^{m} c_i e^{-k_{\sigma(i)}(\lambda_i + j\omega_i)} \right| \\
&\leq m! \max_{\sigma \in S_m} \left| \prod_{i=1}^{m} c_i e^{-k_{\sigma(i)}(\lambda_i + j\omega_i)} \right| \\
&= m! \left| \prod_{i=1}^{m} c_i \right| \max_{\sigma \in S_m} \left| \prod_{i=1}^{m} e^{-k_{\sigma(i)}\lambda_i} \right| \\
&= m! \left| \prod_{i=1}^{m} c_i \right| \prod_{i=1}^{m} e^{-k_i \lambda_i} (\because \textit{Lemma } 7.4) \\
&\lesssim \prod_{i=1}^{m} e^{-k_i \lambda_i}
\end{aligned} \tag{7.5}$$

where c_i are ith component of $\mathbf{C}$, S_m is the set of all permutations on $\{1, \cdots, m\}$, and $sgn(\sigma)$ is $+1$ if σ is an even permutation -1 otherwise. Therefore, the lemma is true for a diagonal $\mathbf{A_c}$.

To extend to a general Jordan matrix $\mathbf{A_c}$, consider a matrix $\mathbf{A'_c}$ which is obtained by erasing the off-diagonal elements of $\mathbf{A_c}$. Then, we can easily see the ratio between the elements of $\begin{bmatrix} \mathbf{C}e^{-k_1 \mathbf{A_c}} \\ \vdots \\ \mathbf{C}e^{-k_m \mathbf{A_c}} \end{bmatrix}$ and the corresponding elements of $\begin{bmatrix} \mathbf{C}e^{-k_1 \mathbf{A'_c}} \\ \vdots \\ \mathbf{C}e^{-k_m \mathbf{A'_c}} \end{bmatrix}$ is a polynomial whose degree is less than m. Therefore, by repeating the steps of (7.5) we can easily obtain

$$\left| \det \left(\begin{bmatrix} \mathbf{C}e^{-k_1 \mathbf{A_c}} \\ \mathbf{C}e^{-k_2 \mathbf{A_c}} \\ \vdots \\ \mathbf{C}e^{-k_m \mathbf{A_c}} \end{bmatrix} \right) \right| \lesssim (1 + k_m^{m^2}) \prod_{i=1}^{m} e^{-k_i \lambda_i},$$

which finishes the proof. □

The next lemma upper bounds the norm of the inverse of the observability Gramian, given the lower bound on the observability Gramian determinant. Therefore, we can reduce the matrix inverse problem to the matrix determinant problem.

Lemma 7.6. *Consider* $\mathbf{A_c}$ *and* $\mathbf{C}$ *given as* (7.3) *and* (7.4). *Let* λ_i *be the real part of* (i,i) *element of* $\mathbf{A_c}$. *Then, there exists a positive polynomial* $p(k)$ *such that for all* $\epsilon > 0$ *and* $0 \leq k_1 \leq \cdots \leq k_m$, *if*

$$\left| \det \left(\begin{bmatrix} \mathbf{C}e^{-k_1\mathbf{A_c}} \\ \mathbf{C}e^{-k_2\mathbf{A_c}} \\ \vdots \\ \mathbf{C}e^{-k_m\mathbf{A_c}} \end{bmatrix} \right) \right| \geq \epsilon \prod_{1 \leq i \leq m} e^{-k_i \lambda_i}$$

then

$$\left| \begin{bmatrix} \mathbf{C}e^{-k_1\mathbf{A_c}} \\ \mathbf{C}e^{-k_2\mathbf{A_c}} \\ \vdots \\ \mathbf{C}e^{-k_m\mathbf{A_c}} \end{bmatrix}^{-1} \right|_{max} \leq \frac{p(k_m)}{\epsilon} e^{\lambda_1 k_m}.$$

Proof. Let $\mathbf{O_{i,j}}$ be the matrix obtained by removing the ith row and jth column of $\begin{bmatrix} \mathbf{C}e^{-k_1\mathbf{A_c}} \\ \mathbf{C}e^{-k_2\mathbf{A_c}} \\ \vdots \\ \mathbf{C}e^{-k_m\mathbf{A_c}} \end{bmatrix}$.
Let $\mathbf{A_c}(\mathbf{j})$ be the $(m-1) \times (m-1)$ matrix that we can obtain by removing the jth row and column of $\mathbf{A_c}$, and $\mathbf{C}(\mathbf{j})$ be the row vector that we can obtain by removing the jth element of $\mathbf{C}$.

First, let's consider the case when $\mathbf{A_c}$ is a diagonal matrix. In this case, using properties of diagonal matrices we can easily check that $\mathbf{O_{i,j}} = \begin{bmatrix} \mathbf{C}(\mathbf{j})e^{-k_1\mathbf{A_c}(\mathbf{j})} \\ \vdots \\ \mathbf{C}(\mathbf{j})e^{-k_{i-1}\mathbf{A_c}(\mathbf{j})} \\ \mathbf{C}(\mathbf{j})e^{-k_{i+1}\mathbf{A_c}(\mathbf{j})} \\ \vdots \\ \mathbf{C}(\mathbf{j})e^{-k_m\mathbf{A_c}(\mathbf{j})} \end{bmatrix}$.

In other words, $\mathbf{O_{i,j}}$ are also the observability Gramian of $(\mathbf{A_c}(\mathbf{j}), \mathbf{C}(\mathbf{j}))$. Let $C_{i,j}$ be the (i,j)th cofactor of $\begin{bmatrix} \mathbf{C}e^{-k_1\mathbf{A_c}} \\ \mathbf{C}e^{-k_2\mathbf{A_c}} \\ \vdots \\ \mathbf{C}e^{-k_m\mathbf{A_c}} \end{bmatrix}$. Since $C_{i,j}$ is the determinant of $\mathbf{O_{i,j}}$, we can apply Lemma 7.5 to conclude that there exists a positive polynomial $p_{i,j}$ such that

$$|C_{i,j}| \leq \begin{cases} p_{i,j}(k_m) \left(\prod_{l=1}^{j-1} e^{-\lambda_l k_l} \right) \cdot \left(\prod_{l=j}^{i-1} e^{-\lambda_{l+1} k_l} \right) \cdot \left(\prod_{l=i+1}^{m} e^{-\lambda_l k_l} \right) & \text{if } i \geq j \\ p_{i,j}(k_m) \left(\prod_{l=1}^{i-1} e^{-\lambda_l k_l} \right) \cdot \left(\prod_{l=i}^{j-1} e^{-\lambda_l k_{l+1}} \right) \cdot \left(\prod_{l=j+1}^{m} e^{-\lambda_l k_l} \right) & \text{if } i \leq j \end{cases} \tag{7.6}$$

Then, let's consider the case when $\mathbf{A_c}$ is a general Jordan form matrix. Compared to the case of a diagonal matrix $\mathbf{A_c}$, the elements of $\mathbf{O_{i,j}}$ only differ by polynomials on k_i in ratio.

Therefore, by the same argument of the proof of Lemma 7.5, we can still find a positive polynomial $p_{i,j}$ satisfying (7.6).

Moreover, since $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 0$ and $0 \leq k_1 \leq k_2 \leq \cdots \leq k_m$, we have

$$\left(\prod_{l=1}^{j-1} e^{-\lambda_l k_l}\right) \cdot \left(\prod_{l=j}^{i-1} e^{-\lambda_{l+1} k_l}\right) \cdot \left(\prod_{l=i+1}^{m} e^{-\lambda_l k_l}\right) \leq \prod_{i=2}^{m} e^{-\lambda_i k_{i-1}},$$

$$\left(\prod_{l=1}^{i-1} e^{-\lambda_l k_l}\right) \cdot \left(\prod_{l=i}^{j-1} e^{-\lambda_l k_{l+1}}\right) \cdot \left(\prod_{l=j+1}^{m} e^{-\lambda_l k_l}\right) \leq \prod_{i=2}^{m} e^{-\lambda_i k_{i-1}}.$$

Therefore, we can further bound the cofactor as follows:

$$|C_{i,j}| \leq \max_{i,j} p_{i,j}(k_m) \prod_{i=2}^{m} e^{-\lambda_i k_{i-1}}.$$

Then, we have

$$\begin{aligned}
\left\| \begin{bmatrix} \mathbf{C}e^{-k_1 \mathbf{A_c}} \\ \mathbf{C}e^{-k_2 \mathbf{A_c}} \\ \vdots \\ \mathbf{C}e^{-k_m \mathbf{A_c}} \end{bmatrix}^{-1} \right\|_{max} &= \frac{\max_{i,j} |C_{i,j}|}{\left| \det\left(\begin{bmatrix} \mathbf{C}e^{-k_1 \mathbf{A_c}} \\ \mathbf{C}e^{-k_2 \mathbf{A_c}} \\ \vdots \\ \mathbf{C}e^{-k_m \mathbf{A_c}} \end{bmatrix} \right) \right|} \leq \frac{\max_{i,j} |C_{i,j}|}{\epsilon \prod_{1 \leq i \leq m} e^{-k_i \lambda_i}} \\
&\leq \frac{\max_{i,j} p_{i,j}(k_m) \prod_{i=2}^{m} e^{-\lambda_i k_{i-1}}}{\epsilon \prod_{1 \leq i \leq m} e^{-k_i \lambda_i}} \\
&= \frac{\max_{i,j} p_{i,j}(k_m)}{\epsilon} e^{\lambda_1 k_1} \prod_{i=2}^{m} e^{\lambda_i (k_i - k_{i-1})} \\
&\leq \frac{\max_{i,j} p_{i,j}(k_m)}{\epsilon} e^{\lambda_1 k_1} \prod_{i=2}^{m} e^{\lambda_1 (k_i - k_{i-1})} (\because \lambda_1 \geq \lambda_i \geq 0, k_i - k_{i-1} \geq 0) \\
&= \frac{\max_{i,j} p_{i,j}(k_m)}{\epsilon} e^{\lambda_1 k_m} \\
&\leq \frac{\sum_{i,j} p_{i,j}(k_m)}{\epsilon} e^{\lambda_1 k_m}
\end{aligned}$$

Therefore, the lemma is true. □

Now, the question is reduced to whether the observability Gramian determinant is large enough. We will find a sufficient condition for the determinant to be large in terms of a simpler analytic function. For this, we first need the following lemma that basically asserts that polynomials increase slower than exponentials.

Lemma 7.7. *For any given polynomial $f(x)$, $\lambda > 0$ and $\epsilon > 0$, there exists $a > 0$ such that*

$$|f(k + x)| \leq \epsilon e^{\lambda \cdot x}$$

for all $x \geq a(\log(k + 1) + 1)$ and $k \geq 0$.

Proof. Let the order of $f(x)$ be p. Then, there exists $c > 0$ such that for all $x \geq 0$,

$$|f(x)| \leq c(1 + x^{p+1}).$$

If we consider $\frac{1}{\lambda}\log\frac{c}{\epsilon} + \frac{1}{\lambda}\log(1 + (2x)^{p+1})$ and x, the former grows logarithmically in x while the later grows linearly on x. Therefore, we can find $t > 0$ such that

$$\frac{1}{\lambda}\log\frac{c}{\epsilon} + \frac{1}{\lambda}\log(1 + (2x)^{p+1}) \leq x$$

for all $x \geq t$. We can also finde $a > 0$ such that $a(\log(k+1)+1) \geq \max\left\{\frac{1}{\lambda}\log\frac{c}{\epsilon} + \frac{1}{\lambda}\log(1 + (2k)^{p+1}), t\right\}$ for all $k \geq 0$.

To check the condition, $|f(k + x)| \leq \epsilon e^{\lambda \cdot x}$, we divide into two cases.

(a) When $x \leq k$,

$|f(k + x)|$ is bounded as follows:

$$\begin{aligned}
|f(k + x)| &\leq c\left(1 + (k + x)^{p+1}\right) \\
&\leq c\left(1 + (2k)^{p+1}\right) \\
&= \epsilon e^{\lambda(\frac{1}{\lambda}\log\frac{c}{\epsilon} + \frac{1}{\lambda}\log(1+(2k)^{p+1}))} \\
&\leq \epsilon e^{\lambda \cdot x}
\end{aligned}$$

where the last inequality comes from $\frac{1}{\lambda}\log\frac{c}{\epsilon} + \frac{1}{\lambda}\log(1 + (2k)^{p+1}) \leq x$.

(b) When $x > k$,

Since $t \leq x$, $\frac{1}{\lambda}\log\frac{c}{\epsilon} + \frac{1}{\lambda}\log(1 + (2x)^{p+1}) \leq x$. Then, we can bound $|f(k + x)|$ as follows:

$$\begin{aligned}
|f(k + x)| &\leq c\left(1 + (k + x)^{p+1}\right) \\
&\leq c\left(1 + (2x)^{p+1}\right) \\
&= \epsilon e^{\lambda(\frac{1}{\lambda}\log\frac{c}{\epsilon} + \frac{1}{\lambda}\log(1+(2x)^{p+1}))} \\
&\leq \epsilon e^{\lambda \cdot x}.
\end{aligned}$$

Therefore, the lemma is proved. □

Now, we give a sufficient condition to guarantee that the determinant of the observability Gramian is large enough.

Lemma 7.8. *Let* $\mathbf{A_c}$ *and* $\mathbf{C}$ *be given as* (7.3) *and* (7.4). *Let* $a_{i,j}$ *and* $C_{i,j}$ *be the* (i, j) *element and cofactor of* $\begin{bmatrix} \mathbf{C}e^{-k_1\mathbf{A_c}} \\ \mathbf{C}e^{-k_2\mathbf{A_c}} \\ \vdots \\ \mathbf{C}e^{-k_m\mathbf{A_c}} \end{bmatrix}$ *respectively. Then there exist* $g_\epsilon(k) : \mathbb{R}^+ \to \mathbb{R}^+$ *and* $a \in \mathbb{R}^+$ *such that for all*

$\epsilon > 0$ and $k_1, \cdots, k_m$ satisfying

$$(i)\ 0 \leq k_1 < k_2 < \cdots < k_m$$
$$(ii)\ k_m - k_{m-1} \geq g_\epsilon(k_{m-1})$$
$$(iii)\ g_\epsilon(k) \leq a(1 + \log(k+1))$$
$$(iv)\ |\sum_{m-m_\mu+1 \leq i \leq m} a_{m,i} C_{m,i}| \geq \epsilon \prod_{1 \leq i \leq m} e^{-k_i \lambda_i}$$

the following inequality holds:

$$\left| \det \left(\begin{bmatrix} \mathbf{C} e^{-k_1 \mathbf{A_c}} \\ \mathbf{C} e^{-k_2 \mathbf{A_c}} \\ \vdots \\ \mathbf{C} e^{-k_m \mathbf{A_c}} \end{bmatrix} \right) \right| \geq \frac{1}{2} \epsilon \prod_{1 \leq i \leq m} e^{-k_i \lambda_i}.$$

Proof. First of all, because $\mathbf{A_c}$ is in Jordan form, it is well known that the elements of $e^{-k\mathbf{A_c}}$ take a specific form [17]. Thus, we can prove that for all $a_{i,j}$ there exists a polynomial $p_{i,j}(k)$ such that $a_{i,j} = p_{i,j}(k_i) e^{-k_i(\lambda_j + j\omega_j)}$. Then, we can find $p(k)$ in the form of $a(1 + k^b)$ $(a > 0)$ such that $p(k) \geq \max_{i,j} |p_{i,j}(k)|$ for all $k \geq 0$. Denote $\lambda' := \lambda_{\mu-1,1} - \lambda_{\mu,1} > 0$.

Let S_m be the set of all permutations on $\{1, \cdots, m\}$, and $sgn(\sigma)$ be $+1$ if σ is an even permutation -1 otherwise. Then, we have

$$\left|\det\left(\begin{bmatrix}\mathbf{C}e^{-k_1\mathbf{A_c}}\\ \mathbf{C}e^{-k_2\mathbf{A_c}}\\ \vdots\\ \mathbf{C}e^{-k_m\mathbf{A_c}}\end{bmatrix}\right)\right| = \left|\sum_{1\le i\le m} a_{m,i}C_{m,i}\right| = \left|\sum_{\sigma\in S_m} sgn(\sigma)\prod_{i=1}^{m} a_{i,\sigma(i)}\right|$$

$$\ge \left|\sum_{m-m_\mu+1\le i\le m} a_{m,i}C_{m,i}\right| - \left|\sum_{1\le i\le m-m_\mu} a_{m,i}C_{m,i}\right|$$

$$\ge \epsilon\prod_{1\le i\le m} e^{-k_i\lambda_i} - \left|\sum_{1\le i\le m-m_\mu} a_{m,i}C_{m,i}\right| (\because \textit{Assumption (iv)})$$

$$= \epsilon\prod_{1\le i\le m} e^{-k_i\lambda_i} - \left|\sum_{\sigma\in S_m, 1\le\sigma(m)\le m-m_\mu} sgn(\sigma)\prod_{i=1}^{m} a_{i,\sigma(i)}\right|$$

$$\ge \epsilon\prod_{1\le i\le m} e^{-k_i\lambda_i} - \sum_{\sigma\in S_m, 1\le\sigma(m)\le m-m_\mu}\left|\prod_{i=1}^{m} a_{i,\sigma(i)}\right|$$

$$= \epsilon\prod_{1\le i\le m} e^{-k_i\lambda_i} - \sum_{\sigma\in S_m, 1\le\sigma(m)\le m-m_\mu}\left|\prod_{i=1}^{m} p_{i,\sigma(i)}(k_i)e^{-k_i(\lambda_{\sigma(i)}+j\omega_{\sigma(i)})}\right|$$

$$\ge \epsilon\prod_{1\le i\le m} e^{-k_i\lambda_i} - \sum_{\sigma\in S_m, 1\le\sigma(m)\le m-m_\mu}\left(e^{(\lambda_m-\lambda_{\sigma(m)})(k_m-k_{\sigma^{-1}(m)})}\cdot\prod_{i=1}^{m} p(k_i)e^{-k_i\lambda_i}\right) (\because \textit{Lemma } 7.4)$$

$$\ge \prod_{1\le i\le m} e^{-k_i\lambda_i}\left(\epsilon - \sum_{\sigma\in S_m, 1\le\sigma(m)\le m-m_\mu} p(k_m)^m e^{(\lambda_m-\lambda_{\sigma(m)})(k_m-k_{\sigma^{-1}(m)})}\right) (\because p(k) \text{ is an increasing function.})$$

$$\ge \prod_{1\le i\le m} e^{-k_i\lambda_i}\left(\epsilon - \sum_{\sigma\in S_m, 1\le\sigma(m)\le m-m_\mu} p(k_m)^m e^{-\lambda'(k_m-k_{m-1})}\right) (\because \lambda_{\sigma(m)}-\lambda_m \ge \lambda_{\mu-1,1}-\lambda_{\mu,1}=\lambda')$$

$$\ge \prod_{1\le i\le m} e^{-k_i\lambda_i}\left(\epsilon - m!p(k_m)^m e^{-\lambda'(k_m-k_{m-1})}\right)$$

Since $m!p(x)^m$ is a polynomial in x, by Lemma 7.7 there exists $g_\epsilon(k):\mathbb{R}^+\to\mathbb{R}^+$ such that
(i) $g_\epsilon(k)\lesssim \log(k+1)+1$
(ii) $|m!p(k+x)^m|\le \frac{\epsilon}{2}e^{\lambda'\cdot x}$ for all $x\ge g_\epsilon(k)$ and $k\ge 0$.
Therefore, for all k_m such that $k_m-k_{m-1}\ge g_\epsilon(k_{m-1})$,

$$\left|\det\left(\begin{bmatrix}\mathbf{C}e^{-k_1\mathbf{A_c}}\\ \mathbf{C}e^{-k_2\mathbf{A_c}}\\ \vdots\\ \mathbf{C}e^{-k_m\mathbf{A_c}}\end{bmatrix}\right)\right| \ge \prod_{1\le i\le m} e^{-k_i\lambda_i}\left(\epsilon-\frac{\epsilon}{2}e^{\lambda'\cdot(k_m-k_{m-1})}e^{-\lambda'\cdot(k_m-k_{m-1})}\right) \ge \frac{\epsilon}{2}\prod_{1\le i\le m} e^{-k_i\lambda_i}.$$

Thus, the lemma is proved. □

7.3 Uniform Convergence of a Set of Analytic Functions (Continuous-Time Systems)

We will prove that after introducing nonuniform sampling, the determinant of the observability Gramian will become large enough regardless of the erasure pattern. Since the determinant of the observability Gramian is an analytic function, to prove that the observability Gramian is large enough it is enough prove that a set of specific analytic functions are large enough. To this end, we will prove a set of analytic functions are uniformly away from 0.

First, we prove that an analytic function can become zero only on sets of zero Lebesgue-measure, as long as the function is not zero for all values. The intuition for the lemma is that analytic functions can be locally determined by that Taylor expansions. Thus, if an analytic function is zero for any open interval with non-zero Lebesgue-measure, it is identically zero.

Lemma 7.9. *For a given nonnegative integer p and distinct positive reals $\omega_{i,1}, \omega_{i,2}, \cdots, \omega_{i,\nu_i}$, define*

$$f(x) := \sum_{i=0}^{p} x^i \left(\sum_{j=1}^{\nu_i} a_{R,i,j} \cos(\omega_{i,j} x) + a_{I,i,j} \sin(\omega_{i,j} x) \right)$$

where at least one coefficient among $a_{R,i,j}, a_{I,i,j}$ is non-zero. Let X be a uniform random variable in $[0, T]$ ($T > 0$). Then, for all $h \in \mathbb{R}$, the following is true:

$$\mathbb{P}\{|f(X) - h| < \epsilon\} \rightarrow 0 \text{ as } \epsilon \downarrow 0.$$

Proof. First, notice that $f(x) - h$ is an analytic function. It is well-known that if an analytic function $f(x) - h$ is not identically zero, the set $\{x \in [0, T] : f(x) - h = 0\}$ is an isolated set [53], which is countable. Therefore, $\mathbb{P}\{|f(X) - h| = 0\} = 0$. Moreover, $\mathbb{P}\{|f(X) - h| < \epsilon\} \leq \mathbb{P}\{|f(X) - h| \leq \epsilon\}$, which is a cumulative distribution function. Since cumulative distribution functions are right-continuous, $\lim_{\epsilon \downarrow 0} \mathbb{P}\{|f(X) - h| < \epsilon\} \leq \lim_{\epsilon \downarrow 0} \mathbb{P}\{|f(X) - h| \leq \epsilon\} = \mathbb{P}\{|f(X) - h| = 0\} = 0$.

Thus, the proof reduces to proving $f(x) - h$ is not zero for all x. Let i^* be the largest i such that either $a_{R,i,j}$ or $a_{I,i,j}$ is non-zero.

(i) When $i^* = 0$,

In this case, there are no polynomial terns and only sinusoidal terms exist. Let's compute the energy of $f(x) - h$ in interval $[s, s + r]$ and prove that $f(x) - h$ is not identically zero for all s

as long as r is large enough.

$$\int_s^{s+r}\left(\sum_{j=1}^{\nu_{i^*}}(a_{R,i^*,j}\cos(\omega_{i^*,j}x)+a_{I,i^*,j}\sin(\omega_{i^*,j}x))-h\right)^2 dx$$

$$=\int_s^{s+r}\sum_{j=1}^{\nu_{i^*}}\left(a_{R,i^*,j}^2\cos^2(\omega_{i^*,j}x)+a_{I,i^*,j}^2\sin^2(\omega_{i^*,j}x)\right)+h^2+2\sum_{i\le j}a_{R,i^*,i}a_{I,i^*,j}\cos(\omega_{i^*,i}x)\sin(\omega_{i^*,j}x)$$

$$+2\sum_{i<j}a_{R,i^*,i}a_{R,i^*,j}\cos(\omega_{i^*,i}x)\cos(\omega_{i^*,j}x)+2\sum_{i<j}a_{I,i^*,i}a_{I,i^*,j}\sin(\omega_{i^*,i}x)\sin(\omega_{i^*,j}x)$$

$$-2\sum_{j=1}^{\nu_{i^*}}(a_{R,i^*,j}\cos(\omega_{i^*,j}x)+a_{I,i^*,j}\sin(\omega_{i^*,j}x))h\ dx$$

$$=\int_s^{s+r}\sum_{j=1}^{\nu_{i^*}}\left(a_{R,i^*,j}^2\frac{1+\cos 2\omega_{i^*,j}x}{2}+a_{I,i^*,j}^2\frac{1-\cos 2\omega_{i^*,j}x}{2}\right)\ dx$$

$$+\int_s^{s+r}\sum_{i\le j}a_{R,i^*,i}a_{I,i^*,j}\left(\sin\left(\left(\omega_{i^*,j}+\omega_{i^*,j}\right)x\right)-\sin\left(\left(\omega_{i^*,j}-\omega_{i^*,j}\right)x\right)\right)\ dx$$

$$+\int_s^{s+r}\sum_{i<j}a_{R,i^*,i}a_{R,i^*,j}\left(\cos\left(\left(\omega_{i^*,j}-\omega_{i^*,j}\right)x\right)+\cos\left(\left(\omega_{i^*,j}+\omega_{i^*,j}\right)x\right)\right)\ dx$$

$$+\int_s^{s+r}\sum_{i<j}a_{I,i^*,i}a_{I,i^*,j}\left(\cos\left(\left(\omega_{i^*,j}-\omega_{i^*,j}\right)x\right)-\cos\left(\left(\omega_{i^*,j}+\omega_{i^*,j}\right)x\right)\right)\ dx$$

$$-\int_s^{s+r}2\sum_{j=1}^{\nu_{i^*}}(a_{R,i^*,j}\cos(\omega_{i^*,j}x)+a_{I,i^*,j}\sin(\omega_{i^*,j}x))h\ dx. \tag{7.7}$$

Therefore, as r increases, the first term in (7.7) arbitrarily increases regardless of s, while the remaining terms in (7.7) are sinusoidal and so bounded. Thus, $f(x)-h$ is not identically zero for all s when r is large enough. Thus, there exist $\delta>0$ and $r>0$ such that for all s, $|f(x)-x|\ge\delta$ holds for some $x\in[s,s+r]$.

(ii) When $i^*\ge 1$,

In this case, we have polynomial terms and we will prove that the term with the highest degree will dominate the reaming terms. By the argument of (i), we can find $\delta>0$ and $r>0$ such that for all $s\ge 0$ we can find $x\in[s,s+r]$ satisfying

$$|f(x)-h|\ge\delta x^{i^*}-\sum_{i=0}^{i^*-1}\left(\sum_{j=1}^{\nu_i}|a_{R,i,j}|+|a_{I,i,j}|\right)x^i-|h|.$$

Since we can choose s arbitrarily large, $|f(x)-h|$ has to be greater than 0 for some x. Thus, $f(x)-h$ is not identically zero.

Therefore, the lemma is true. □

To prove uniform convergence, we need the following Dini's theorem which says that for compact sets, pointwise convergence implies uniform convergence. The intuition behind this theorem is as follows: since we can find a finite open cover for a compact set, we can convert the uniform convergence of an infinite number of functions to the uniform convergence of only finitely many functions when the domain is compact. The uniform convergence of a finite number of functions immediately follows from pointwise convergence.

Theorem 7.1 (Dini's Theorem)**.** *[35, p. 81] If $\{f_n\}$ is a sequence of functions defined on a set A and converging on A to a function f, and if*
(i) the convergence is monotonic,
(ii) f_n is continuous on A, $n = 1, 2, \cdots$
(iii) f is continuous on A,
(iv) A is compact,
then the convergence is uniform on A.

Proof. See [35, p. 81] for the proof. □

Now, using the pointwise convergence of Lemma 7.9 and Dini's theorem, we can prove the uniform convergence of the relevant functions over a set of parameters.

Lemma 7.10. *Let p, $\nu_0, \cdots, \nu_p$ be nonnegative integers with $\nu_p > 0$. Suppose γ and Γ are strictly positive reals such that $\gamma \leq \Gamma$. For each $0 \leq i \leq p$, $\omega_{i,1}, \omega_{i,2}, \cdots, \omega_{i,\nu_i}$ are distinct reals. Let X be a uniform random variable on $[0, T]$ for some $T > 0$. Then, for all m, n such that $0 \leq m \leq p$ and $1 \leq n \leq \nu_m$, we have the following inequality:*

$$\sup_{|a_{m,n}| \geq \gamma, \forall i,j, |a_{i,j}| \leq \Gamma} \mathbb{P}\left\{ \left| \sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} a_{i,j} e^{j\omega_{i,j} X} \right) \right| < \epsilon \right\} \to 0 \text{ as } \epsilon \downarrow 0$$

where $a_{i,j}$ are taken from $\mathbb{C}$.

Proof. The purpose of this proof is reducing the lemma to Dini's theorem (Theorem 7.1).

First, we will assume the $w_{i,j}$ are positive without loss of generality. To justify this, let $\omega_{min} = \min\{\min_{i,j} \omega_{i,j}, 0\} - \delta$ for some $\delta > 0$. Then,

$$\begin{aligned} &\sup_{|a_{m,n}| \geq \gamma, |a_{i,j}| \leq \Gamma} \mathbb{P}\left\{ \left| \sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} a_{i,j} e^{j\omega_{i,j} X} \right) \right| < \epsilon \right\} \\ &= \sup_{|a_{m,n}| \geq \gamma, |a_{i,j}| \leq \Gamma} \mathbb{P}\left\{ \left| \sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} a_{i,j} e^{j(\omega_{i,j} - \omega_{min}) X} \right) \right| < \epsilon \right\}. \end{aligned}$$

Here, for each i, $\omega_{i,1} - \omega_{min}, \omega_{i,2} - \omega_{min}, \cdots, \omega_{i,\nu_i} - \omega_{min}$ are distinct and strictly positive. Therefore, without loss of generality, we can assume that for each i, $\omega_{i,1}, \omega_{i,2}, \cdots, \omega_{i,\nu_i}$ are distinct and strictly positive.

Let $a_{i,j} = a_{R,i,j} - ja_{I,i,j}$ where $a_{R,i,j}$ and $a_{I,i,j}$ are real. Since $|a_{m,n}| \geq \gamma$, at least one of $|a_{R,m,n}|$ or $|a_{I,m,n}|$ should be greater than $\frac{\gamma}{\sqrt{2}}$. First, consider the case when $|a_{R,m,n}| \geq \frac{\gamma}{\sqrt{2}}$. It is sufficient to prove that the real part of $\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} a_{i,j} e^{j\omega_{i,j}X}\right)$ satisfies the lemma, i.e.

$$\sup_{a_{R,m,n}\geq\frac{\gamma}{\sqrt{2}},|a_{R,i,j}|\leq\Gamma,|a_{I,i,j}|\leq\Gamma} \mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} a_{R,i,j}\cos(\omega_{i,j}X) + a_{I,i,j}\sin(\omega_{i,j}X)\right)\right| < \epsilon\right\} \to 0 \text{ as } \epsilon \downarrow 0.$$

Here, we take the supremum over $a_{R,m,n} \geq \frac{\gamma}{\sqrt{2}}$ instead of the supremum over $|a_{R,m,n}| \geq \frac{\gamma}{\sqrt{2}}$ by symmetry.

Now, we apply Dini's theorem 7.1 and prove the claim.

Fix a positive sequence ϵ_i such that $\epsilon_i \downarrow 0$ as $i \to \infty$. Define a sequence of functions $\{f_i\}$ as

$$f_i(a_{R,1,1}, a_{I,1,1}, \cdots, a_{I,p,\nu_p}) := \mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} a_{R,i,j}\cos(\omega_{i,j}X) + a_{I,i,j}\sin(\omega_{i,j}X)\right)\right| < \epsilon_i\right\}$$

where the domain A of the functions is $A := \{(a_{R,1,1}, a_{I,1,1}, \cdots, a_{I,p,\nu_p}) : a_{R,m,n} \geq \frac{\gamma}{\sqrt{2}}, |a_{R,i,j}| \leq \Gamma, |a_{I,i,j}| \leq \Gamma\}$. Let $f(a_{R,1,1}, a_{I,1,1}, \cdots, a_{I,p,\nu_p})$ be the identically zero function. Then, we will prove that $\{f_i\}$ converges to $f = 0$ uniformly on A by checking the conditions of Theorem 7.1.

• f_i point-wisely converges to f:

Since $a_{R,m,n} \geq \frac{\gamma}{\sqrt{2}}$, $\sum_{i=0}^{p} x^i \left(\sum_{j=1}^{\nu_i} a_{R,i,j}\cos(\omega_{i,j}x) + a_{I,i,j}\sin(\omega_{i,j}x)\right)$ satisfies the assumptions of Lemma 7.9. Thus, for all h

$$\mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} a_{R,i,j}\cos(\omega_{i,j}X) + a_{I,i,j}\sin(\omega_{i,j}X)\right) - h\right| < \epsilon\right\} \to 0 \text{ as } \epsilon \downarrow 0. \quad (7.8)$$

Therefore, by selecting $h = 0$, $f_i(a_{R,1,1}, a_{I,1,1}, \cdots, a_{I,p,\nu_p})$ converges to $f = 0$ for all $a_{R,1,1}$, $a_{I,1,1}$, $\cdots$, a_{I,p,ν_p} in A.

• Convergence is monotone: Since ϵ_i monotonically converge to 0, f_i is also a monotonically decreasing function sequence. Thus, the convergence is monotone.

• f_n is continuous on A: For continuity (does not have to be uniformly continuous), we will prove that for given $a_{R,1,1}, a_{I,1,1}, \cdots, a_{I,p,\nu_p}$ and for all $\sigma > 0$, there exists $\delta(\sigma) > 0$ such that $|f_i(a_{R,1,1} + \nabla a_{R,1,1}, a_{I,1,1} + \nabla a_{I,1,1}, \cdots, a_{I,p,\nu_p} + \nabla a_{I,p,\nu_p}) - f_i(a_{R,1,1}, a_{I,1,1}, \cdots, a_{I,p,\nu_p})| < \sigma$ for all $|\nabla a_{R,i,j}| < \delta(\sigma)$ and $|\nabla a_{I,i,j}| < \delta(\sigma)$.

By (7.8), we can find $\delta'(\sigma)$ for all σ such that

$$\mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} a_{R,i,j}\cos(\omega_{i,j}X) + a_{I,i,j}\sin(\omega_{i,j}X)\right) - \epsilon_i\right| < \delta'(\sigma)\right\} < \frac{\sigma}{2} \text{ and}$$
$$\mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} a_{R,i,j}\cos(\omega_{i,j}X) + a_{I,i,j}\sin(\omega_{i,j}X)\right) - (-\epsilon_i)\right| < \delta'(\sigma)\right\} < \frac{\sigma}{2}.$$

Denote $\delta(\sigma) := \frac{\min\left(\frac{1}{T^p},1\right)}{2\sum_{i=0}^{p}\nu_i}\delta'(\sigma)$. Then, for all $|\nabla a_{R,i,j}| < \delta(\sigma)$ and $|\nabla a_{I,i,j}| < \delta(\sigma)$, the following inequality is true.

$$
\begin{aligned}
&\mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i\left(\sum_{j=1}^{\nu_i}(a_{R,i,j}+\nabla a_{R,i,j})\cos(\omega_{i,j}X)+(a_{I,i,j}+\nabla a_{R,i,j})\sin(\omega_{i,j}X)\right)\right|<\epsilon_i\right\}\\
&\geq \mathbb{P}\Bigg\{\left|\sum_{i=0}^{p} X^i\left(\sum_{j=1}^{\nu_i}a_{R,i,j}\cos(\omega_{i,j}X)+a_{I,i,j}\sin(\omega_{i,j}X)\right)\right|<\\
&\epsilon_i-\left|\sum_{i=0}^{p} X^i\left(\sum_{j=1}^{\nu_i}\nabla a_{R,i,j}\cos(\omega_{i,j}X)+\nabla a_{I,i,j}\sin(\omega_{i,j}X)\right)\right|\Bigg\}\\
&\geq \mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i\left(\sum_{j=1}^{\nu_i}a_{R,i,j}\cos(\omega_{i,j}X)+a_{I,i,j}\sin(\omega_{i,j}X)\right)\right|<\epsilon_i-\delta'(\sigma)\right\} \qquad (7.9)\\
&= \mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i\left(\sum_{j=1}^{\nu_i}a_{R,i,j}\cos(\omega_{i,j}X)+a_{I,i,j}\sin(\omega_{i,j}X)\right)\right|<\epsilon_i\right\}\\
&\quad-\mathbb{P}\left\{\epsilon_i-\delta'(\sigma)\leq\left|\sum_{i=0}^{p} X^i\left(\sum_{j=1}^{\nu_i}a_{R,i,j}\cos(\omega_{i,j}X)+a_{I,i,j}\sin(\omega_{i,j}X)\right)\right|<\epsilon_i\right\}\\
&\geq \mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i\left(\sum_{j=1}^{\nu_i}a_{R,i,j}\cos(\omega_{i,j}X)+a_{I,i,j}\sin(\omega_{i,j}X)\right)\right|<\epsilon_i\right\}\\
&\quad-\mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i\left(\sum_{j=1}^{\nu_i}a_{R,i,j}\cos(\omega_{i,j}X)+a_{I,i,j}\sin(\omega_{i,j}X)\right)-\epsilon_i\right|<\delta'(\sigma)\right\}\\
&\quad-\mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i\left(\sum_{j=1}^{\nu_i}a_{R,i,j}\cos(\omega_{i,j}X)+a_{I,i,j}\sin(\omega_{i,j}X)\right)-(-\epsilon_i)\right|<\delta'(\sigma)\right\}\\
&> \mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i\left(\sum_{j=1}^{\nu_i}a_{R,i,j}\cos(\omega_{i,j}X)+a_{I,i,j}\sin(\omega_{i,j}X)\right)\right|<\epsilon_i\right\}-\sigma.
\end{aligned}
$$

Here, (7.9) can be shown as follows:

$$
\begin{aligned}
&\left|\sum_{i=0}^{p} X^i\left(\sum_{j=1}^{\nu_i}\nabla a_{R,i,j}\cos(\omega_{i,j}X)+\nabla a_{I,i,j}\sin(\omega_{i,j}X)\right)\right|\\
&\leq \sum_{i=0}^{p}|X^i|\sum_{j=1}^{\nu_i}(|\nabla a_{R,i,j}|+|\nabla a_{I,i,j}|)\\
&\leq \max(T^p,1)2\nu_i\delta(\sigma)(\because 0\leq X\leq T \text{ w.p. } 1)\\
&= \delta'(\sigma)(\because \text{definition of } \delta(\sigma))
\end{aligned}
$$

Therefore, by the definition of f_i we have

$$
\begin{aligned}
&f_i(a_{R,1,1}+\nabla a_{R,1,1},a_{I,1,1}+\nabla a_{I,1,1},\cdots,a_{I,p,\nu_p}+\nabla a_{I,p,\nu_p})\\
&-f_i(a_{R,1,1},a_{I,1,1},\cdots,a_{I,p,\nu_p})>-\sigma. \qquad (7.10)
\end{aligned}
$$

Likewise, we can prove that

$$\begin{aligned}
&\mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} (a_{R,i,j}+\nabla a_{R,i,j})\cos(\omega_{i,j}X) + (a_{I,i,j}+\nabla a_{R,i,j})\sin(\omega_{i,j}X)\right)\right| < \epsilon_i\right\}\\
&\le \mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} a_{R,i,j}\cos(\omega_{i,j}X) + a_{I,i,j}\sin(\omega_{i,j}X)\right)\right| < \epsilon_i + \delta'(\sigma)\right\}\\
&(\because \text{The same step as (7.9)})\\
&= \mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} a_{R,i,j}\cos(\omega_{i,j}X) + a_{I,i,j}\sin(\omega_{i,j}X)\right)\right| < \epsilon_i\right\}\\
&\quad + \mathbb{P}\left\{\epsilon_i \le \left|\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} a_{R,i,j}\cos(\omega_{i,j}X) + a_{I,i,j}\sin(\omega_{i,j}X)\right)\right| < \epsilon_i + \delta'(\sigma)\right\}\\
&\le \mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} a_{R,i,j}\cos(\omega_{i,j}X) + a_{I,i,j}\sin(\omega_{i,j}X)\right)\right| < \epsilon_i\right\}\\
&\quad + \mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} a_{R,i,j}\cos(\omega_{i,j}X) + a_{I,i,j}\sin(\omega_{i,j}X)\right) - \epsilon_i\right| < \delta'(\sigma)\right\}\\
&\quad + \mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} a_{R,i,j}\cos(\omega_{i,j}X) + a_{I,i,j}\sin(\omega_{i,j}X)\right) - (-\epsilon_i)\right| < \delta'(\sigma)\right\}\\
&< \mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} a_{R,i,j}\cos(\omega_{i,j}X) + a_{I,i,j}\sin(\omega_{i,j}X)\right)\right| < \epsilon_i\right\} + \sigma
\end{aligned}$$

which implies

$$\begin{aligned}
&f_i(a_{R,1,1}+\nabla a_{R,1,1}, a_{I,1,1}+\nabla a_{I,1,1}, \cdots, a_{I,p,\nu_p}+\nabla a_{I,p,\nu_p})\\
&- f_i(a_{R,1,1}, a_{I,1,1}, \cdots, a_{I,p,\nu_p}) < \sigma. \qquad (7.11)
\end{aligned}$$

By (7.10) and (7.11),

$$\left|f_i(a_{R,1,1}+\nabla a_{R,1,1}, a_{I,1,1}+\nabla a_{I,1,1}, \cdots, a_{I,p,\nu_p}+\nabla a_{I,p,\nu_p}) - f_i(a_{R,1,1}, a_{I,1,1}, \cdots, a_{I,p,\nu_p})\right| < \sigma.$$

Therefore, $f_i(a_{R,1,1}, a_{I,1,1}, \cdots, a_{I,p,\nu_p})$ is continuous.

- f is continuous on A: f is obviously continuous, since f is identically zero.
- A is compact: A is compact since it is closed and bounded.

Thus, by Dini's theorem 7.1, the convergence is uniform on A, which finishes the proof for the case of $|a_{R,m,n}| \ge \frac{\gamma}{\sqrt{2}}$. The proof for the case of $|a_{I,m,n}| \ge \frac{\gamma}{\sqrt{2}}$ follows in an identical manner. Since there are only two cases, the function

$$g_i(a_{1,1}, \cdots, a_{p,\nu_p}) := \mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} a_{i,j} e^{j\omega_{i,j}X}\right)\right| < \epsilon_i\right\}$$

converges uniformly on $\{a_{i,j} : |a_{m,n}| \geq \gamma, |a_{i,j}| \leq \Gamma\}$. This finishes the proof of the lemma. □

In Lemma 7.10, we have a boundedness condition on the coefficients ($|a_{i,j}| \leq \Gamma$) to guarantee compactness. However, we can easily notice the functions only get larger as $a_{i,j}$ increases. Therefore, we can prove that Lemma 7.10 still holds without the boundedness condition.

Lemma 7.11. *Let p be a nonnegative integer and $\nu_0, \cdots, \nu_p$ be also nonnegative integers with $\nu_p > 0$. γ is a strictly positive real. For each $0 \leq i \leq p$, $\omega_{i,1}, \omega_{i,2}, \cdots, \omega_{i,\nu_i}$ are distinct reals. Let X be a uniform random variable on $[0, T]$ for some $T > 0$. Then, for all m, n such that $0 \leq m \leq p$ and $1 \leq n \leq \nu_m$, we have the following inequality:*

$$\sup_{|a_{m,n}| \geq \gamma} \mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} a_{i,j} e^{j\omega_{i,j}X}\right)\right| < \epsilon\right\} \to 0 \text{ as } \epsilon \downarrow 0$$

where $a_{i,j}$ are taken from $\mathbb{C}$.

Proof. Denote $\nu := \sum_{i=0}^{p} \nu_i$. The proof is by strong induction on ν.

(i) When $\nu = 1$.

$$\sup_{|a_{p,1}| \geq \gamma} \mathbb{P}\left\{\left|a_{p,1} X^p e^{j\omega_{p,1}X}\right| < \epsilon\right\} \tag{7.12}$$

$$= \sup_{|a_{p,1}| \geq \gamma} \mathbb{P}\left\{\left|\frac{\gamma}{|a_{p,1}|} a_{p,1} X^p e^{j\omega_{p,1}X}\right| < \frac{\gamma}{|a_{p,1}|}\epsilon\right\}$$

$$\leq \sup_{|a'_{p,1}| = \gamma} \mathbb{P}\left\{\left|a'_{p,1} X^p e^{j\omega_{p,1}X}\right| < \epsilon\right\} \left(\because \frac{\gamma}{|a_{p,1}|} \leq 1\right) \tag{7.13}$$

By lemma 7.10, (7.13) converges to 0 as $\epsilon \downarrow 0$. Thus, (7.12) converges to 0 as $\epsilon \downarrow 0$.

(ii) As an induction hypothesis, we assume the lemma is true for $\nu = 1, \cdots, n-1$ and prove that the lemma still holds for $\nu = n$. We will prove this by dividing into two cases: (a) When all $a_{i,j}$ are not much bigger than $a_{m,n}$. In this case, the claim reduces to Lemma 7.10. (b) When there is an $a_{m',n'}$ which is much bigger than $a_{m,n}$. In this case, we can ignore the term associated with $a_{m,n}$ and reduce the number of terms in the functions. Thus, either way the claim reduces to the induction hypothesis.

To prove the lemma for $\nu = n$, it is enough to show that for a fixed γ and every $\delta > 0$, there exists $\epsilon(\delta) > 0$ such that

$$\sup_{|a_{m,n}| \geq \gamma} \mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} a_{i,j} e^{j\omega_{i,j}X}\right)\right| < \epsilon(\delta)\right\} < \delta.$$

By the induction hypothesis for all $(m', n') \neq (m, n)$ we can find $\epsilon_{m',n'}(\delta) > 0$ such that

$$\sup_{a_{m,n}=0, |a_{m',n'}| \geq \gamma} \mathbb{P}\left\{\left|\sum_{i=0}^{p} X^p \left(\sum_{j=1}^{\nu_i} a_{i,j} e^{j\omega_{i,j}X}\right)\right| < \epsilon_{m',n'}(\delta)\right\} < \delta. \tag{7.14}$$

We choose $\kappa(\delta)$ as $\min\left\{\min_{(m',n')\neq(m,n)}\left\{\frac{\epsilon_{m',n'}(\delta)}{2\gamma T^m}\right\},1\right\}$. By Lemma 7.10, there exists $\epsilon'(\delta) > 0$ such that

$$\sup_{|a_{m,n}|=\gamma, a_{i,j}\leq\frac{\gamma}{\kappa(\delta)}} \mathbb{P}\left\{\left|\sum_{i=0}^{p} X^p \left(\sum_{j=1}^{\nu_i} a_{i,j}e^{j\omega_{i,j}X}\right)\right| < \epsilon'(\delta)\right\} < \delta. \tag{7.15}$$

Denote $\epsilon(\delta) := \min\left\{\epsilon'(\delta), \min_{(m',n')\neq(m,n)}\left\{\frac{\epsilon_{m',n'}(\delta)}{2}\right\}\right\}$. Then, we have

$$\begin{aligned}
&\sup_{|a_{m,n}|\geq\gamma} \mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} a_{i,j}e^{j\omega_{i,j}X}\right)\right| < \epsilon(\delta)\right\}\\
&= \max\{\sup_{|a_{m,n}|\geq\gamma, \frac{|a_{i,j}|}{|a_{m,n}|}\leq\frac{1}{\kappa(\delta)}} \mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} a_{i,j}e^{j\omega_{i,j}X}\right)\right| < \epsilon(\delta)\right\}, && (7.16)\\
&\max_{(m',n')\neq(m,n)} \sup_{|a_{m,n}|\geq\gamma, \frac{|a_{m',n'}|}{|a_{m,n}|}\geq\frac{1}{\kappa(\delta)}} \mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} a_{i,j}e^{j\omega_{i,j}X}\right)\right| < \epsilon(\delta)\right\} && (7.17)\\
&\}. && (7.18)
\end{aligned}$$

• When the $a_{i,j}$ are not too bigger than $a_{m,n}$: Let's bound the first term in (7.16). Set $a'_{i,j} := \frac{\gamma}{|a_{m,n}|}a_{i,j}$. Then, (7.16) is upper bounded as follows:

$$\begin{aligned}
&\sup_{|a_{m,n}|\geq\gamma, \frac{|a_{i,j}|}{|a_{m,n}|}\leq\frac{1}{\kappa(\delta)}} \mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} a_{i,j}e^{j\omega_{i,j}X}\right)\right| < \epsilon(\delta)\right\}\\
&= \sup_{|a_{m,n}|\geq\gamma, \frac{|a_{i,j}|}{|a_{m,n}|}\leq\frac{1}{\kappa(\delta)}} \mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} \frac{\gamma}{|a_{m,n}|}a_{i,j}e^{j\omega_{i,j}X}\right)\right| < \frac{\gamma}{|a_{m,n}|}\epsilon(\delta)\right\}\\
&= \sup_{|a'_{m,n}|=\gamma, |a'_{i,j}|\leq\frac{\gamma}{\kappa(\delta)}} \mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} a'_{i,j}e^{j\omega_{i,j}X}\right)\right| < \frac{\gamma}{|a_{m,n}|}\epsilon(\delta)\right\}\\
&\leq \sup_{|a'_{m,n}|=\gamma, |a'_{i,j}|\leq\frac{\gamma}{\kappa(\delta)}} \mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} a'_{i,j}e^{j\omega_{i,j}X}\right)\right| < \epsilon(\delta)\right\} (\because \frac{\gamma}{|a_{m,n}|}\leq 1)\\
&\leq \sup_{|a'_{m,n}|=\gamma, |a'_{i,j}|\leq\frac{\gamma}{\kappa(\delta)}} \mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} a'_{i,j}e^{j\omega_{i,j}X}\right)\right| < \epsilon'(\delta)\right\} (\because \text{definition of } \epsilon(\delta))\\
&< \delta (\because (7.15)) && (7.19)
\end{aligned}$$

• When $a_{m',n'}$ is much bigger than $a_{m,n}$: Let's bound the second term in (7.17). For given

m', n', set $a''_{i,j} := \frac{\gamma}{|a_{m',n'}|} a_{i,j}$. Then, (7.17) is upper bounded by

$$\max_{(m',n')\neq(m,n)} \sup_{|a_{m,n}|\geq\gamma, \frac{|a_{m',n'}|}{|a_{m,n}|}\geq\frac{1}{\kappa(\delta)}} \mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} a_{i,j} e^{j\omega_{i,j}X}\right)\right| < \epsilon(\delta)\right\}$$

$$= \max_{(m',n')\neq(m,n)} \sup_{|a_{m,n}|\geq\gamma, \frac{|a_{m',n'}|}{|a_{m,n}|}\geq\frac{1}{\kappa(\delta)}} \mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} \frac{\gamma}{|a_{m',n'}|} a_{i,j} e^{j\omega_{i,j}X}\right)\right| < \frac{\gamma}{|a_{m',n'}|}\epsilon(\delta)\right\}$$

$$\leq \max_{(m',n')\neq(m,n)} \sup_{|a_{m,n}|\geq\gamma, \frac{|a_{m',n'}|}{|a_{m,n}|}\geq\frac{1}{\kappa(\delta)}} \mathbb{P}\Bigg\{\left|\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} \frac{\gamma}{|a_{m',n'}|} a_{i,j} e^{j\omega_{i,j}X}\right) - X^m \frac{\gamma}{|a_{m',n'}|} a_{m,n} e^{j\omega_{m,n}X}\right|$$

$$< \max_{(m',n')\neq(m,n)} \frac{\gamma}{|a_{m',n'}|}\epsilon(\delta) + \frac{\gamma}{|a_{m',n'}|}|a_{m,n}|T^m\Bigg\}$$

$$\leq \max_{(m',n')\neq(m,n)} \sup_{|a_{m,n}|\geq\gamma, \frac{|a_{m',n'}|}{|a_{m,n}|}\geq\frac{1}{\kappa(\delta)}} \mathbb{P}\Bigg\{\left|\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} \frac{\gamma}{|a_{m',n'}|} a_{i,j} e^{j\omega_{i,j}X}\right) - X^m \frac{\gamma}{|a_{m',n'}|} a_{m,n} e^{j\omega_{m,n}X}\right|$$

$$< \epsilon_{m',n'}(\delta)\Bigg\} \tag{7.20}$$

$$\leq \max_{(m',n')\neq(m,n)} \sup_{a''_{m,n}=0, |a''_{m',n'}|=\gamma} \mathbb{P}\left\{\left|\sum_{i=0}^{p-1} X^i \left(\sum_{j=1}^{\nu_i} a''_{i,j} e^{j\omega_{i,j}X}\right)\right| < \epsilon_{m',n'}(\delta)\right\}$$

$(\because$ By definition, $a''_{m',n'} = \dfrac{\gamma}{|a_{m',n'}|} a_{m',n'})$

$$< \delta (\because (7.14)) \tag{7.21}$$

Here, (7.20) can be derived as follows: First, we have

$$\begin{aligned}
1 &\geq \kappa(\delta) \quad (\because \text{Definition of } \kappa(\delta)) \\
&\geq \frac{\gamma \cdot \kappa(\delta)}{|a_{m,n}|} \quad (\because |a_{m,n}| \geq \gamma) \\
&\geq \frac{\gamma}{|a_{m',n'}|}. \quad (\because \frac{|a_{m',n'}|}{|a_{m,n}|} \geq \frac{1}{\kappa(\delta)})
\end{aligned} \tag{7.22}$$

We also have

$$\begin{aligned}
\frac{\gamma}{|a_{m',n'}|}|a_{m,n}|T^m &\leq \gamma \cdot \kappa(\delta) T^m \quad (\because \frac{|a_{m',n'}|}{|a_{m,n}|} \geq \frac{1}{\kappa(\delta)}) \\
&\leq \gamma \frac{\epsilon_{m',n'}(\delta)}{2\gamma T^m} T^m \quad (\because \text{By definition, } \kappa(\delta) \leq \frac{\epsilon_{m',n'}(\delta)}{2\gamma T^m}) \\
&= \frac{\epsilon_{m',n'}(\delta)}{2}.
\end{aligned} \tag{7.23}$$

Therefore,

$$\begin{aligned}
\frac{\gamma}{|a_{m',n'}|}\epsilon(\delta) + \frac{\gamma}{|a_{m',n'}|}|a_{m,n}|T^m &\leq \epsilon(\delta) + \frac{\epsilon_{m',n'}(\delta)}{2} \quad (\because (7.22), (7.23)) \\
&\leq \epsilon_{m',n'}(\delta). \quad (\because \text{By definition, } \epsilon(\delta) \leq \frac{\epsilon_{m',n'}(\delta)}{2})
\end{aligned}$$

Therefore, (7.20) is true.

By plugging (7.19) and (7.21) into (7.18), we get

$$\sup_{|a_{m,n}|\geq\gamma} \mathbb{P}\left\{\left|\sum_{i=0}^{p} X^i \left(\sum_{j=1}^{\nu_i} a_{i,j} e^{j\omega_{i,j}X}\right)\right| < \epsilon(\delta)\right\} < \delta,$$

which finishes the proof. □

7.4 Proof of Lemma 2.2

In this section, we will merge the properties about the observability Gramian shown in Section 7.2 with the uniform convergence of Section 7.3, and prove Lemma 2.2 of page 56.

We first prove the following lemma which tells us that the determinant of the observability Gramian is large with high probability under a cofactor condition on the Gramian.

Lemma 7.12. *Let $\mathbf{A_c}$ and $\mathbf{C}$ be given as* (7.3) *and* (7.4)*. Let $a_{i,j}$ and $C_{i,j}$ be the (i,j) element and cofactor of* $\begin{bmatrix} \mathbf{C}e^{-(k_1 I+t_1)\mathbf{A_c}} \\ \vdots \\ \mathbf{C}e^{-(k_{m-1} I+t_{m-1})\mathbf{A_c}} \\ \mathbf{C}e^{-(k_m I+t)\mathbf{A_c}} \end{bmatrix}$ *respectively, where t is a random variable which is uniformly distributed on $[0,T]$ and I is the sampling interval defined in* (2.25)*. Then, there exist $a \in \mathbb{R}^+$ and a family of increasing functions $\{g_\epsilon(\cdot) : \epsilon > 0, g_\epsilon : \mathbb{R}^+ \to \mathbb{R}^+\}$ satisfying:*
(i) For all $\epsilon > 0$, $k_1 < k_2 < \cdots < k_{m-1}$, $0 \leq t_i \leq T$ if $|C_{m,m}| > \epsilon \prod_{1\leq i\leq m-1} e^{-k_i I\cdot\lambda_i}$ the following is true:

$$\sup_{k_m\in\mathbb{Z}, k_m-k_{m-1}\geq g_\epsilon(k_{m-1})} \mathbb{P}\left\{\left|\det\left(\begin{bmatrix} \mathbf{C}e^{-(k_1 I+t_1)\mathbf{A_c}} \\ \vdots \\ \mathbf{C}e^{-(k_{m-1} I+t_{m-1})\mathbf{A_c}} \\ \mathbf{C}e^{-(k_m I+t)\mathbf{A_c}} \end{bmatrix}\right)\right| < \epsilon^2 \prod_{1\leq i\leq m} e^{-k_i I\cdot\lambda_i}\right\} \to 0 \text{ as } \epsilon \downarrow 0$$

(ii) For all $\epsilon > 0$, $g_\epsilon(k) \leq a(1+\log(k+1))$.

Proof. Let $\epsilon' = 2\epsilon^2 \prod_{1\leq i\leq m} e^{\lambda_i T}$. Define $a'_{i,j}$, $C'_{i,j}$ as the (i,j) element and cofactor of $\begin{bmatrix} \mathbf{C}e^{-\kappa_1\mathbf{A_c}} \\ \vdots \\ \mathbf{C}e^{-\kappa_m\mathbf{A_c}} \end{bmatrix}$.

Then, by Lemma 7.8, we can find a function $g'_{\epsilon'}(k)$ such that for all $0 \leq \kappa_1 < \kappa_2 < \cdots < \kappa_m$ satisfying:
(i') $\kappa_m - \kappa_{m-1} \geq g'_{\epsilon'}(\kappa_{m-1})$
(ii') $g'_{\epsilon'}(\kappa) \lesssim 1 + \log(\kappa+1)$
(iii') $|\sum_{m-m_\mu+1\leq i\leq m} a'_{m,i} C'_{m,i}| \geq \epsilon' \prod_{1\leq i\leq m} e^{-\kappa_i\lambda_i}$

the following inequality holds:

$$\left|\det\left(\begin{bmatrix}\mathbf{C}e^{-\kappa_1\mathbf{A_c}}\\ \vdots \\ \mathbf{C}e^{-\kappa_m\mathbf{A_c}}\end{bmatrix}\right)\right| \geq \frac{1}{2}\epsilon' \prod_{1\leq i\leq m} e^{-\kappa_i\lambda_i}.$$

Let's use t_m and t interchangeably. These are in $[0, T]$ with probability one. Ideally, we want to plug $k_iI + t_i$ into κ_i. However, even though the sequence $k_1, \cdots, k_m$ is sorted, the sequence $k_1I + t_1, \cdots, k_mI + t_m$ may not be sorted. Therefore, we define $k_{(1)}I + t_{(1)}, \cdots, k_{(m)}I + t_{(m)}$ as the result of sorting $k_1I + t_1, \cdots, k_mI + t_m$. Then, we can see this sorted sequence has the following property.

Claim 7.1. *Consider two sequences,* $\alpha_1, \alpha_2, \cdots, \alpha_n$ *and* $\beta_1, \beta_2, \cdots, \beta_n$ *where* $\alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_n$ *and* $\beta_i \in [0, T]$ $(T > 0)$*. Let* $\alpha_{(1)} + \beta_{(1)}, \alpha_{(2)} + \beta_{(2)}, \cdots, \alpha_{(n)} + \beta_{(n)}$ *be the ascending ordered set of* $\alpha_1 + \beta_1, \alpha_2 + \beta_2, \cdots, \alpha_n + \beta_n$*. In other words,*

Then, for all $i \in \{1, \cdots, n\}$*, we have*

$$0 \leq \alpha_{(i)} + \beta_{(i)} - \alpha_i \leq T.$$

Proof. We will prove this by contradiction. Let's say there exists i such that

$$\alpha_{(i)} + \beta_{(i)} - \alpha_i < 0.$$

Then, we have

$$\alpha_{(i)} + \beta_{(i)} < \alpha_i \leq \alpha_{i+1} \leq \cdots \leq \alpha_n.$$

Since $\beta_1, \cdots, \beta_n \geq 0$, we can conclude $\alpha_{(i)} + \beta_{(i)} < \alpha_i + \beta_i$, $\cdots$, $\alpha_{(i)} + \beta_{(i)} < \alpha_n + \beta_n$. Thus, in the sequence $\alpha_1 + \beta_1, \cdots, \alpha_n + \beta_n$, there exist $n - i + 1$ elements which are larger than $\alpha_{(i)} + \beta(i)$. This contradicts the fact that $\alpha_{(i)} + \beta_{(i)}$ is ith largest element among $\alpha_1 + \beta_1, \cdots, \alpha_n + \beta_n$.

Likewise, let's say there exists i such that

$$\alpha_{(i)} + \beta_{(i)} - \alpha_i > T.$$

Then, we have

$$\alpha_{(i)} + \beta_{(i)} > \alpha_i + T \leq \alpha_{i-1} + T \leq \alpha_1 + T.$$

Since $\beta_1, \cdots, \beta_n \leq T$, we can conclude $\alpha_{(i)} + \beta_{(i)} > \alpha_i + \beta_i$, $\cdots$, $\alpha_{(i)} + \beta_{(i)} > \alpha_1 + \beta_1$. Thus, in the sequence $\alpha_1 + \beta_1, \cdots, \alpha_n + \beta_n$, there exist i elements which are smaller than $\alpha_{(i)} + \beta(i)$. This contradicts the fact that $\alpha_{(i)} + \beta_{(i)}$ is ith smallest element among $\alpha_1 + \beta_1, \cdots, \alpha_n + \beta_n$. ☐

Therefore, by the claim, we have

$$\prod_{1\le i\le m} e^{-\lambda_i T} \prod_{1\le i\le m} e^{-k_i I\cdot\lambda_i} \le \prod_{1\le i\le m} e^{-(k_{(i)}I+t_{(i)})\lambda_i} \le \prod_{1\le i\le m} e^{-k_i I\cdot\lambda_i}. \tag{7.24}$$

Finally, we can plug $k_{(i)}I + t_{(i)}$ into κ_i to conclude the following statement. For all $0 \le k_1 < \cdots < k_m$, $0 \le t_i \le T$, $0 \le t \le T$ such that[1]

(i") $k_m - k_{m-1} \ge g''_{\epsilon'}(k_{m-1})$

(ii") $g''_{\epsilon'}(k) \lesssim 1 + \log(k+1)$

(iii") $\left|\sum_{m-m_\mu+1\le i\le m} a_{m,i}C_{m,i}\right| \ge \epsilon' \prod_{1\le i\le m} e^{-k_i I\cdot\lambda_i} \overset{(A)}{\ge} \epsilon' \prod_{1\le i\le m} e^{-(k_{(i)}I+t_{(i)})\lambda_i}$

the following inequality holds:

$$\left|\det\left(\begin{bmatrix} \mathbf{C}e^{-(k_1 I+t_1)\mathbf{A_c}} \\ \vdots \\ \mathbf{C}e^{-(k_{m-1}I+t_{m-1})\mathbf{A_c}} \\ \mathbf{C}e^{-(k_m I+t)\mathbf{A_c}} \end{bmatrix}\right)\right| \ge \frac{1}{2}\epsilon' \prod_{1\le i\le m} e^{-(k_{(i)}I+t_{(i)})\lambda_i} \overset{(B)}{\ge} \frac{1}{2}\epsilon' \prod_{1\le i\le m} e^{-\lambda_i T} \prod_{1\le i\le m} e^{-k_i I\cdot\lambda_i}$$
$$\overset{(C)}{=} \epsilon^2 \prod_{1\le i\le m} e^{-k_i I\cdot\lambda_i}.$$

Here, (A) and (B) always hold by (7.24). (C) follows from the definition of ϵ'.

Let $g_\epsilon(k)$ be $g''_{\epsilon'}(k)$. Then, we can easily check such $g_\epsilon(k)$ satisfies condition (ii) of the lemma. Let's show that such $g_\epsilon(k)$ also satisfies condition (i) of the lemma.

$$\sup_{k_m\in\mathbb{Z}, k_m-k_{m-1}\ge g_\epsilon(k_{m-1})} \mathbb{P}\left\{\left|\det\left(\begin{bmatrix} \mathbf{C}e^{-(k_1 I+t_1)\mathbf{A_c}} \\ \vdots \\ \mathbf{C}e^{-(k_{m-1}I+t_{m-1})\mathbf{A_c}} \\ \mathbf{C}e^{-(k_m I+t)\mathbf{A_c}} \end{bmatrix}\right)\right| < \epsilon^2 \prod_{1\le i\le m} e^{-k_i I\cdot\lambda_i}\right\}$$
$$\le \sup_{k_m\in\mathbb{Z}, k_m-k_{m-1}\ge g_\epsilon(k_{m-1})} \mathbb{P}\left\{\left|\sum_{m-m_\mu+1\le i\le m} C_{m,i}a_{m,i}\right| < 2\epsilon^2 \prod_{1\le i\le m} e^{\lambda_i T}\cdot \prod_{1\le i\le m} e^{-k_i I\cdot\lambda_i}\right\}$$
$$= \sup_{k_m\in\mathbb{Z}, k_m-k_{m-1}\ge g_\epsilon(k_{m-1})} \mathbb{P}\left\{\left|\sum_{m-m_\mu+1\le i\le m} \frac{C_{m,i}}{\epsilon\prod_{1\le i\le m-1} e^{-k_i I\cdot\lambda_i}} \frac{a_{m,i}}{e^{-(k_m I+t)\lambda_m}}\right| < 2\epsilon\cdot e^{\lambda_m t} \prod_{1\le i\le m} e^{\lambda_i T}\right\}$$
$$\le \sup_{|b_m|\ge 1} \mathbb{P}\left\{\left|\sum_{m-m_\mu+1\le i\le m} b_i \frac{a_{m,i}}{e^{-(k_m I+t)\lambda_m}}\right| < 2\epsilon\cdot e^{\lambda_m T} \prod_{1\le i\le m} e^{\lambda_i T}\right\}. \tag{7.25}$$

where the last inequality comes from assumption (ii), $|C_{m,m}| > \epsilon \prod_{1\le i\le m-1} e^{-k_i I\cdot\lambda_i}$, and the fact that $t \in [0,T]$ with probability one.

Now, it is enough to prove that (7.25) converges to 0 as $\epsilon \downarrow 0$. To this end, let's study $a_{m,i}$ which are the elements of the observability gramian. Let the $\mathbf{C}_{\mu,\nu_\mu}$ defined in (7.4)

[1] Here, we select $g''_{\epsilon'}(k)$ large enough so that when $k_m - k_{m-1} \ge g''_{\epsilon'}(k_{m-1})$, we always have $k_m I + t \ge k_{m-1}I + t_{m-1}$, i.e. $k_m I + t$ becomes the largest.

be $\begin{bmatrix} c'_1 & \cdots & c'_{m_{\mu,\nu_\mu}} \end{bmatrix}$. Then, we have

$$e^{-(k_m I+t)\mathbf{A}_{\mu,\nu_\mu}}$$
$$= \begin{bmatrix} e^{-(k_m I+t)(\lambda_{\mu,\nu_\mu}+j\omega_{\mu,\nu_\mu})} & -(k_m I+t)e^{-(k_m I+t)(\lambda_{\mu,\nu_\mu}+j\omega_{\mu,\nu_\mu})} & \cdots & \frac{(-1)^{m_{\mu,\nu_\mu}-1}(k_m I+t)^{m_{\mu,\nu_\mu}-1}}{(m_{\mu,\nu_\mu}-1)!}e^{-(k_m I+t)(\lambda_{\mu,\nu_\mu}+j\omega_{\mu,\nu_\mu})} \\ 0 & e^{-(k_m I+t)(\lambda_{\mu,\nu_\mu}+j\omega_{\mu,\nu_\mu})} & \cdots & \frac{(-1)^{m_{\mu,\nu_\mu}-2}(k_m I+t)^{m_{\mu,\nu_\mu}-2}}{(m_{\mu,\nu_\mu}-2)!}e^{-(k_m I+t)(\lambda_{\mu,\nu_\mu}+j\omega_{\mu,\nu_\mu})} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{-(k_m I+t)(\lambda_{\mu,\nu_\mu}+j\omega_{\mu,\nu_\mu})} \end{bmatrix}.$$

Thus, we can see that

$$a_{m,m} = \sum_{1\le i\le m_{\mu,\nu_\mu}} c'_i \frac{(-1)^{m_{\mu,\nu_\mu}-i}(k_m I+t)^{m_{\mu,\nu_\mu}-i}}{(m_{\mu,\nu_\mu}-i)!} e^{-(k_m I+t)(\lambda_m+j\omega_{\mu,\nu_\mu})}.$$

Therefore,

$$\frac{a_{m,m}}{e^{-(k_m I+t)\lambda_m}} = \sum_{1\le i\le m_{\mu,\nu_\mu}} c'_i \frac{(-1)^{m_{\mu,\nu_\mu}-i}(k_m I+t)^{m_{\mu,\nu_\mu}-i}}{(m_{\mu,\nu_\mu}-i)!} e^{-(k_m I+t)(j\omega_{\mu,\nu_\mu})}.$$

Moreover, when $a_{m,i}$ is considered as a function of t, the $t^{m_{\mu,\nu_\mu}-1}e^{-j\omega_{\mu,\nu_\mu}t}$ term only shows up in $\frac{a_{m,m}}{e^{-(k_m I+t)\lambda_m}}$ among $\frac{a_{m,m-m_\mu+1}}{e^{-(k_m I+t)\lambda_m}}, \cdots, \frac{a_{m,m}}{e^{-(k_m I+t)\lambda_m}}$, and the coefficient is $c'_1 \frac{(-1)^{m_{\mu,\nu_\mu}-1}}{(m_{\mu,\nu_\mu}-1)!} e^{-j\omega_{\mu,\nu_\mu} k_m I}$. Since we know $|b_m| \ge 1$ in (7.25), by defining $c' := \frac{|c'_1|}{(m_{\mu,\nu_\mu}-1)!}$ we can see that the magnitude of the corresponding coefficient is greater or equal to c'. Furthermore, the remaining terms $\frac{a_{m,m-m_\mu+1}}{e^{-(k_m I+t)\lambda_m}}$, $\cdots$, $\frac{a_{m,m-1}}{e^{-(k_m I+t)\lambda_m}}$ only have $e^{-j\omega_{\mu,1}t}$, $\cdots$, $t^{m_{\mu,1}-1}e^{-j\omega_{\mu,1}t}$, $e^{-j\omega_{\mu,2}t}$, $\cdots$, $t^{m_{\mu,2}-1}e^{-j\omega_{\mu,2}t}$, $\cdots$, $e^{-j\omega_{\mu,\nu_\mu}t}$, $\cdots$, $t^{m_{\mu,\nu_\mu}-2}e^{-j\omega_{\mu,\nu_\mu}t}$ when they are considered as functions in t. Thus, using the assumption that $m_{\nu,1} \le \cdots \le m_{\nu,\mu_\nu}$,(7.25) can be upper bounded as follows:

$$(7.25) \le \sup_{|a'_{m_{\mu,\nu_\mu},\nu_\mu}|\ge c'} \mathbb{P}\left\{ \left| \sum_{i=1}^{m_{\mu,\nu_\mu}} t^{i-1} \left(\sum_{j=1}^{\nu_\mu} a'_{i,j} e^{-j\omega_{\mu,j}t} \right) \right| \le 2\epsilon e^{\lambda_m T} \cdot \prod_{1\le i\le m} e^{\lambda_i T} \right\}. \quad (7.26)$$

By Lemma 7.11 (by setting γ as c', (m,n) as $(m_{\mu,\nu_\mu}, \nu_\mu)$, p as m_{μ,ν_μ}, $\nu_0, \cdots, \nu_p$ as ν_μ, $\omega_{0,j}, \cdots, \omega_{p,j}$ as $-\omega_{\mu,j}$, and ϵ as $2\epsilon \prod_{1\le i\le m} e^{\lambda_i T} \cdot e^{\lambda_m T}$), we get

$$\sup_{|a'_{m_{\mu,\nu_\mu},\nu_\mu}|\ge c'} \mathbb{P}\left\{ \left| \sum_{i=1}^{m_{\mu,\nu_\mu}} t^{i-1} \left(\sum_{j=1}^{\nu_\mu} a'_{i,j} e^{-j\omega_{\mu,j}t} \right) \right| \le 2\epsilon e^{\lambda_m T} \cdot \prod_{1\le i\le m} e^{\lambda_i T} \right\} \to 0 \text{ as } \epsilon \downarrow 0. \quad (7.27)$$

Therefore, by (7.25), (7.26), (7.27) we can say that

$$\sup_{k_m\in\mathbb{Z}, k_m-k_{m-1}\ge g_\epsilon(k_{m-1})} \mathbb{P}\left\{ \left| \det\left(\begin{bmatrix} \mathbf{C}e^{-(k_1 I+t_1)\mathbf{A_c}} \\ \vdots \\ \mathbf{C}e^{-(k_{m-1} I+t_{m-1})\mathbf{A_c}} \\ \mathbf{C}e^{-(k_m I+t)\mathbf{A_c}} \end{bmatrix} \right) \right| < \epsilon^2 \prod_{1\le i\le m} e^{-k_i I\cdot\lambda_i} \right\} \to 0 \text{ as } \epsilon \downarrow 0$$

which finishes the proof. $\square$

Based on the previous lemma, we will integrate the properties of p.m.f. tails shown in Section 7.1 with the properties of the observability Gramian discussed in Section 7.2, and prove Lemma 2.2 for the case of a row vector $\mathbf{C}$.

Lemma 7.13. *Let* $\mathbf{A_c}$ *and* $\mathbf{C}$ *be given as* (7.3) *and* (7.4)*. Let* $\beta[n]$ $(n \in \mathbb{Z}^+)$ *be a Bernoulli random process with probability* $1-p_e$ *and* t_n *be i.i.d. random variables which are uniformly distributed on* $[0,T]$ $(T>0)$*. Then, we can find a polynomial* $p(k)$ *and a family of stopping times* $\{S(\epsilon,k) : k \in \mathbb{Z}^+, \epsilon > 0\}$ *such that for all* $\epsilon > 0$*,* $k \in \mathbb{Z}^+$ *there exist* $k \leq k_1 < k_2 < \cdots < k_m \leq S(\epsilon,k)$ *and* $\mathbf{M}$ *satisfying the following conditions:*
(i) $\beta[k_i] = 1$ *for* $1 \leq i \leq m$
(ii) $\mathbf{M}\begin{bmatrix} \mathbf{C}e^{-(k_1 I + t_{k_1})\mathbf{A_c}} \\ \mathbf{C}e^{-(k_2 I + t_{k_2})\mathbf{A_c}} \\ \vdots \\ \mathbf{C}e^{-(k_m I + t_{k_m})\mathbf{A_c}} \end{bmatrix} = \mathbf{I}$
(iii) $|\mathbf{M}|_{max} \leq \frac{p(S(\epsilon,k))}{\epsilon} e^{\lambda_1 S(\epsilon,k) I}$
(iv) $\lim_{\epsilon \downarrow 0} \exp \limsup_{s\to\infty} \sup_{k\in\mathbb{Z}^+} \frac{1}{s} \log \mathbb{P}\{S(\epsilon,k) - k = s\} \leq p_e$*.*

Proof. By Lemma 7.6, instead of conditions (ii) and (iii), it is enough to prove that

$$\left|\det\left(\begin{bmatrix} \mathbf{C}e^{-(k_1 I + t_{k_1})\mathbf{A_c}} \\ \mathbf{C}e^{-(k_2 I + t_{k_2})\mathbf{A_c}} \\ \vdots \\ \mathbf{C}e^{-(k_m I + t_{k_m})\mathbf{A_c}} \end{bmatrix}\right)\right| \geq \epsilon \prod_{1\leq i \leq m} e^{-(k_i I + t_{k_i})\lambda_i}.$$

Furthermore, since $t_i \geq 0$ it is sufficient to prove that

$$\left|\det\left(\begin{bmatrix} \mathbf{C}e^{-(k_1 I + t_{k_1})\mathbf{A_c}} \\ \mathbf{C}e^{-(k_2 I + t_{k_2})\mathbf{A_c}} \\ \vdots \\ \mathbf{C}e^{-(k_m I + t_{k_m})\mathbf{A_c}} \end{bmatrix}\right)\right| \geq \epsilon \prod_{1\leq i \leq m} e^{-k_i I \cdot \lambda_i}.$$

Therefore, it is enough to prove the following claim:

We can find a family of stopping times $\{S(\epsilon,k) : k \in \mathbb{Z}^+, \epsilon > 0\}$ such that for all $\epsilon > 0$ and $k \in \mathbb{Z}^+$ there exist $k \leq k_1 < k_2 < \cdots < k_m \leq S(\epsilon,k)$ satisfying the following condition:
(a) $\beta[k_i] = 1$ for $1 \leq i \leq m$
(b) $\left|\det\left(\begin{bmatrix} \mathbf{C}e^{-(k_1 I + t_{k_1})\mathbf{A_c}} \\ \mathbf{C}e^{-(k_2 I + t_{k_2})\mathbf{A_c}} \\ \vdots \\ \mathbf{C}e^{-(k_m I + t_{k_m})\mathbf{A_c}} \end{bmatrix}\right)\right| \geq \epsilon \prod_{1\leq i \leq m} e^{-k_i I \cdot \lambda_i}$
(c) $\lim_{\epsilon \downarrow 0} \exp \limsup_{s\to\infty} \sup_{k\in\mathbb{Z}^+} \frac{1}{s} \log \mathbb{P}\{S(\epsilon,k) - k = s\} \leq p_e$

We will prove the claim by induction on m, the size of the $\mathbf{A_c}$ matrix.

(i) When $m = 1$,

Since we only have to care about small enough ϵ, assume $\epsilon \le |c_1| e^{-2T\lambda_1}$. Denote $S(\epsilon, k) := \inf\{n \ge k : \beta[n] = 1\}$ and $k_1 = S(\epsilon, k)$. Then, $\beta[k_1] = 1$ and $\left|\det\left(\left[c_1 e^{-(k_1 I + t_{k_1})(\lambda_1 + j\omega_1)}\right]\right)\right| \ge |c_1| e^{-T\lambda_1} e^{-k_1 I \cdot \lambda_1} \ge \epsilon e^{-k_1 I \cdot \lambda_1}$.
Moreover, since $S(\epsilon, k) - k$ is a geometric random variable with probability $1 - p_e$,

$$\exp \limsup_{s \to \infty} \sup_{k \in \mathbb{Z}^+} \log \mathbb{P}\left\{S(\epsilon, k) - k = s\right\} = p_e.$$

Therefore, $S(\epsilon, k)$ satisfies all the conditions of the lemma.

(ii) Now, we assume that the lemma is true for $m-1$ and prove the lemma still holds for m.

First, we will fix $k = 0$, then we will consider general $k \in \mathbb{Z}^+$. We will see that the induction hypothesis corresponds to the cofactor condition of Lemma 7.12, which tells us that the determinant of the observability Gramian is large enough with high probability.

Let $\mathbf{A}'_{\mathbf{c}}$ be the $(m-1) \times (m-1)$ matrix obtained by removing mth row and column of $\mathbf{A}_{\mathbf{c}}$. Likewise, $\mathbf{C}'$ is a $1 \times (m-1)$ vector obtained by removing mth element of $\mathbf{C}$. Then, since $\mathbf{A}_{\mathbf{c}}$ is given in a Jordan form, we can easily check that once we remove the last element from the row vector $\mathbf{C} e^{-(k_i I + t_{k_i})\mathbf{A}_{\mathbf{c}}}$, we get $\mathbf{C}' e^{-(k_i I + t_{k_i})\mathbf{A}'_{\mathbf{c}}}$. Therefore, we can see that

$$\det\left(\begin{bmatrix} \mathbf{C}' e^{-(k_1 I + t_{k_1})\mathbf{A}'_{\mathbf{c}}} \\ \vdots \\ \mathbf{C}' e^{-(k_{m-1} I + t_{k_{m-1}})\mathbf{A}'_{\mathbf{c}}} \end{bmatrix}\right) = cof_{m,m}\left(\begin{bmatrix} \mathbf{C} e^{-(k_1 I + t_{k_1})\mathbf{A}_{\mathbf{c}}} \\ \vdots \\ \mathbf{C} e^{-(k_m I + t_{k_m})\mathbf{A}_{\mathbf{c}}} \end{bmatrix}\right) \tag{7.28}$$

where $cof_{i,j}(\mathbf{A})$ implies the cofactor matrix of $\mathbf{A}$ with respect to (i, j) element.

By the induction hypothesis, there exists a stopping time $S'(\epsilon, 0)$ such that we can find $0 \le k_1 < k_2 < \cdots < k_{m-1} \le S'(\epsilon, 0)$ satisfying:
(a) $\beta[k_i] = 1$ for $1 \le i \le m-1$
(b) $\left|\det\left(\begin{bmatrix} \mathbf{C}' e^{-(k_1 I + t_{k_1})\mathbf{A}'_{\mathbf{c}}} \\ \vdots \\ \mathbf{C}' e^{-(k_{m-1} I + t_{k_{m-1}})\mathbf{A}'_{\mathbf{c}}} \end{bmatrix}\right)\right| \ge \epsilon \prod_{1 \le i \le m-1} e^{-k_i I \cdot \lambda_i}$
(c) $\lim_{\epsilon \downarrow 0} \exp \limsup_{s \to \infty} \frac{1}{s} \log \mathbb{P}\left\{S'(\epsilon, 0) = s\right\} \le p_e$.

Let $\mathcal{F}_i$ be a σ-field generated by $\beta[0], \cdots, \beta[i]$, and $t_0, \cdots, t_i$. Let $g_\epsilon : \mathbb{R}^+ \to \mathbb{R}^+$ be the function of Lemma 7.12. Denote

$$p'(\epsilon) := \operatorname{ess\,sup} \sup_{k_m \in \mathbb{Z}, k_m - S'(\epsilon,0) \ge g_\epsilon(S'(\epsilon,0))} \mathbb{P}_t \left\{ \left| \det\left(\begin{bmatrix} \mathbf{C} e^{-(k_1 I + t_{k_1})\mathbf{A}_{\mathbf{c}}} \\ \vdots \\ \mathbf{C} e^{-(k_{m-1} I + t_{k_{m-1}})\mathbf{A}_{\mathbf{c}}} \\ \mathbf{C} e^{-(k_m I + t)\mathbf{A}_{\mathbf{c}}} \end{bmatrix}\right) \right| < \epsilon^2 \prod_{1 \le i \le m} e^{-k_i I \cdot \lambda_i} | \mathcal{F}_{S'(\epsilon,0)} \right\}. \tag{7.29}$$

Here, given $\mathcal{F}_{S'(\epsilon,0)}$, $k_1, \cdots, k_{m-1}$, $t_{k_1}, \cdots, t_{k_{m-1}}$, $S'(\epsilon, 0)$ are all fixed, we took the supremum over k_m such that $k_m - S'(\epsilon, 0) \geq g_\epsilon(S'(\epsilon, 0))$, and t is a uniform random variable on $[0, T]$ which we computed the probability over.

Since $k_m \geq S'(\epsilon, 0) + g_\epsilon(S'(\epsilon, 0)) \geq k_{m-1} + g_\epsilon(k_{m-1})$ and we have (7.28), (b) implies $cof_{m,m}\left(\begin{bmatrix} \mathbf{C}e^{-(k_1 I + t_{k_1})\mathbf{A_c}} \\ \vdots \\ \mathbf{C}e^{-(k_m I + t_{k_m})\mathbf{A_c}} \end{bmatrix}\right) \geq \epsilon \prod_{1 \leq i \leq m-1} e^{-k_i I \cdot \lambda_i}$. Thus, by Lemma 7.12 we have $\lim_{\epsilon \downarrow 0} p'(\epsilon) = 0$.

Denote $S''(\epsilon, 0) := \lceil S'(\epsilon, 0) + g_\epsilon(S'(\epsilon, 0)) \rceil$. From (ii) of Lemma 7.12 we know $g_\epsilon(k) \lesssim 1 + \log(k+1)$ for all $\epsilon > 0$. Therefore, by (c) and Lemma 7.1 we have

$$\lim_{\epsilon \downarrow 0} \exp \limsup_{s \to \infty} \frac{1}{s} \log \mathbb{P}\{S''(\epsilon, 0) = s\} \leq p_e. \tag{7.30}$$

Denote a stopping time

$$S'''(\epsilon, 0)$$
$$:= \inf \left\{ n \geq S''(\epsilon) : \beta[n] = 1 \text{ and } \left| \det \left(\begin{bmatrix} \mathbf{C}e^{-(k_1 I + t_{k_1})\mathbf{A_c}} \\ \vdots \\ \mathbf{C}e^{-(k_{m-1} I + t_{k_{m-1}})\mathbf{A_c}} \\ \mathbf{C}e^{-(nI + t_n)\mathbf{A_c}} \end{bmatrix} \right) \right| \geq \epsilon^2 e^{-nI \cdot \lambda_m} \prod_{1 \leq i \leq m-1} e^{-k_i I \cdot \lambda_i} \right\}. \tag{7.31}$$

Since $\beta[n]$ and t_n are independent processes, for $S'''(\epsilon, 0) = n$ to hold, $\beta[n] = 1$ and the determinant of (7.31) has to be large enough. By (7.29), we already know the probability for the determinant not being large enough is upper bounded by $p'(\epsilon)$. Therefore, given that $S'''(\epsilon, 0) \geq n$, the probability that $S'''(\epsilon, 0) \neq n$ is upper bounded by $(p_e + (1 - p_e)p'(\epsilon))$ — (erasure) or (not erased but small determinant). Thus, for all $s \in \mathbb{Z}^+$, we have

$$\operatorname{ess\,sup} \mathbb{P}\{S'''(\epsilon, 0) - S''(\epsilon, 0) \geq s | \mathcal{F}_{S''(\epsilon,0)}\} \leq (p_e + (1 - p_e)\, p'(\epsilon))^s.$$

Since we know $\lim_{\epsilon \downarrow 0} p'(\epsilon) = 0$, we have

$$\lim_{\epsilon \downarrow 0} \exp \limsup_{s \to \infty} \operatorname{ess\,sup} \frac{1}{s} \log \mathbb{P}\{S'''(\epsilon, 0) - S''(\epsilon, 0) = s | \mathcal{F}_{S''(\epsilon,0)}\} \leq p_e. \tag{7.32}$$

By applying Lemma 7.2 to (7.30) and (7.32), we can conclude that

$$\lim_{\epsilon \downarrow 0} \exp \limsup_{s \to \infty} \frac{1}{s} \log \mathbb{P}\{S'''(\epsilon, 0) = s\} \leq p_e.$$

Therefore, if we denote $S(\epsilon, 0) := S'''(\epsilon^{\frac{1}{2}}, 0)$, $S(\epsilon, 0)$ satisfies all the conditions of the claim when we fix $k = 0$.

Here, we know $\beta[n]$ is stationary process. Thus, to prove the claim for general $k \in \mathbb{Z}^+$, we can shift the time index by k. Then, we can find a family of stopping times $\{S(\epsilon, k) : k \in \mathbb{Z}^+, \epsilon > 0\}$

such that for all $\epsilon > 0$ and $k \in \mathbb{Z}^+$ there exist $k \leq k_1 < k_2 < \cdots < k_m \leq S(\epsilon, k)$ satisfying the following condition:

(a') $\beta[k_i] = 1$ for $1 \leq i \leq m$

(b') $\left| \det \left(\begin{bmatrix} \mathbf{C} e^{-((k_1-k)I+t_{k_1})\mathbf{A_c}} \\ \mathbf{C} e^{-((k_2-k)I+t_{k_2})\mathbf{A_c}} \\ \vdots \\ \mathbf{C} e^{-((k_m-k)I+t_{k_m})\mathbf{A_c}} \end{bmatrix} \right) \right| \geq \epsilon \prod_{1 \leq i \leq m} e^{-(k_i-k)I \cdot \lambda_i}$

(c') $\lim_{\epsilon \downarrow 0} \exp \limsup_{s \to \infty} \sup_{k \in \mathbb{Z}^+} \frac{1}{s} \log \mathbb{P}\{S(\epsilon, k) - k = s\} \leq p_e$

Here, we can notice that the condition (b') is equivalent to

$$\left| \det \left(\begin{bmatrix} \mathbf{C} e^{-(k_1 I+t_{k_1})\mathbf{A_c}} \\ \mathbf{C} e^{-(k_2 I+t_{k_2})\mathbf{A_c}} \\ \vdots \\ \mathbf{C} e^{-(k_m I+t_{k_m})\mathbf{A_c}} \end{bmatrix} \right) \right| \cdot |\det(e^{kI\mathbf{A_c}})| \geq \epsilon \prod_{1 \leq i \leq m} e^{-(k_i-k)I \cdot \lambda_i}$$

$$(\Leftrightarrow) \left| \det \left(\begin{bmatrix} \mathbf{C} e^{-(k_1 I+t_{k_1})\mathbf{A_c}} \\ \mathbf{C} e^{-(k_2 I+t_{k_2})\mathbf{A_c}} \\ \vdots \\ \mathbf{C} e^{-(k_m I+t_{k_m})\mathbf{A_c}} \end{bmatrix} \right) \right| \geq |\det(e^{kI\mathbf{A_c}})|^{-1} \cdot \epsilon \prod_{1 \leq i \leq m} e^{-(k_i-k)I \cdot \lambda_i} = \epsilon \prod_{1 \leq i \leq m} e^{-k_i I \cdot \lambda_i}$$

Therefore, the claim is true for all $k \in \mathbb{Z}^+$ and the lemma is also true. □

Before we prove Lemma 2.2, we will first prove the following lemma which allows us to merge two Jordan blocks associated with the same eigenvalue into one Jordan block.

Lemma 7.14. *Let* $\mathbf{A}$ *be a Jordan block matrix with an eigenvalue* $\lambda \in \mathbb{C}$ *and a size* $m \in \mathbb{N}$*, i.e.* $\mathbf{A} = \begin{bmatrix} \lambda & 1 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda \end{bmatrix}$. $\mathbf{C}$ *and* $\mathbf{C}'$ *are* $1 \times m$ *matrices such that*

$$\mathbf{C} = \begin{bmatrix} c_1 & c_2 & \cdots & c_m \end{bmatrix}$$
$$\mathbf{C}' = \begin{bmatrix} c'_1 & c'_2 & \cdots & c'_m \end{bmatrix}$$

where $c_i, c'_i \in \mathbb{C}$ *and* $c_1 \neq 0$.

For all $k \in \mathbb{R}$ *and* $m \times 1$ *matrices* $\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$ *and* $\mathbf{X}' = \begin{bmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_m \end{bmatrix}$, *there exists* $\mathbf{T}$ *such that*

$(i)\mathbf{T}$ *is an upper triangular matrix.*

$(ii)\mathbf{C}e^{k\mathbf{A}}\mathbf{X} + \mathbf{C}'e^{k\mathbf{A}}\mathbf{X}' = \mathbf{C}e^{k\mathbf{A}}\left(\mathbf{X} + \mathbf{T}\mathbf{X}'\right)$

Moreover, the diagonal elements of $\mathbf{T}$ *are* $\frac{c_1'}{c_1}$.

Proof. The proof is an induction on m, the size of the $\mathbf{A}$ matrix. The lemma is trivial when $m = 1$. Thus, we can assume the lemma is true for m as an induction hypothesis, and consider $m+1$ as the dimension of $\mathbf{A}$.

$$\mathbf{C}e^{k\mathbf{A}}\mathbf{X} + \mathbf{C}'e^{k\mathbf{A}}\mathbf{X}'$$

$$= \mathbf{C}\begin{bmatrix} e^{k\lambda} & \frac{k}{1!}e^{k\lambda} & \cdots & \frac{k^m}{m!}e^{k\lambda} \\ 0 & e^{k\lambda} & \cdots & \frac{k^{m-1}}{(m-1)!}e^{k\lambda} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{k\lambda} \end{bmatrix}\mathbf{X} + \mathbf{C}'\begin{bmatrix} e^{k\lambda} & \frac{k}{1!}e^{k\lambda} & \cdots & \frac{k^m}{m!}e^{k\lambda} \\ 0 & e^{k\lambda} & \cdots & \frac{k^{m-1}}{(m-1)!}e^{k\lambda} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{k\lambda} \end{bmatrix}\mathbf{X}'$$

$$= \mathbf{C}\begin{bmatrix} e^{k\lambda} & \frac{k}{1!}e^{k\lambda} & \cdots & \frac{k^m}{m!}e^{k\lambda} \\ 0 & e^{k\lambda} & \cdots & \frac{k^{m-1}}{(m-1)!}e^{k\lambda} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{k\lambda} \end{bmatrix}\mathbf{X}$$

$$+ \left(\frac{c_1'}{c_1}\mathbf{C} + \begin{bmatrix} 0 & c_2' - \frac{c_1'}{c_1}c_2 & \cdots & c_{m+1}' - \frac{c_1'}{c_1}c_{m+1} \end{bmatrix}\right)\begin{bmatrix} e^{k\lambda} & \frac{k}{1!}e^{k\lambda} & \cdots & \frac{k^m}{m!}e^{k\lambda} \\ 0 & e^{k\lambda} & \cdots & \frac{k^{m-1}}{(m-1)!}e^{k\lambda} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{k\lambda} \end{bmatrix}\mathbf{X}'$$

$$= \mathbf{C}\begin{bmatrix} e^{k\lambda} & \frac{k}{1!}e^{k\lambda} & \cdots & \frac{k^m}{m!}e^{k\lambda} \\ 0 & e^{k\lambda} & \cdots & \frac{k^{m-1}}{(m-1)!}e^{k\lambda} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{k\lambda} \end{bmatrix}\left(\mathbf{X} + \frac{c_1'}{c_1}\mathbf{X}'\right)$$

$$+ \begin{bmatrix} 0 & c_2' - \frac{c_1'}{c_1}c_2 & \cdots & c_m' - \frac{c_1'}{c_1}c_m \end{bmatrix}\begin{bmatrix} e^{k\lambda} & \frac{k}{1!}e^{k\lambda} & \cdots & \frac{k^m}{m!}e^{k\lambda} \\ 0 & e^{k\lambda} & \cdots & \frac{k^{m-1}}{(m-1)!}e^{k\lambda} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{k\lambda} \end{bmatrix}\mathbf{X}'$$

$$= \mathbf{C}\begin{bmatrix} e^{k\lambda} & \frac{k}{1!}e^{k\lambda} & \cdots & \frac{k^m}{m!}e^{k\lambda} \\ 0 & e^{k\lambda} & \cdots & \frac{k^{m-1}}{(m-1)!}e^{k\lambda} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{k\lambda} \end{bmatrix}\left(\begin{bmatrix} 0 \\ 0 \\ \vdots \\ x_{m+1} + \frac{c_1'}{c_1}x_{m+1}' \end{bmatrix} + \begin{bmatrix} x_1 + \frac{c_1'}{c_1}x_1' \\ x_2 + \frac{c_1'}{c_1}x_2' \\ \vdots \\ 0 \end{bmatrix}\right)$$

$$+ \begin{bmatrix} c_2' - \frac{c_1'}{c_1}c_2 & c_3' - \frac{c_1'}{c_1}c_3 & \cdots & c_{m+1}' - \frac{c_1'}{c_1}c_{m+1} \end{bmatrix}\begin{bmatrix} e^{k\lambda} & \frac{k}{1!}e^{k\lambda} & \cdots & \frac{k^{m-1}}{(m-1)!}e^{k\lambda} \\ 0 & e^{k\lambda} & \cdots & \frac{k^{m-2}}{(m-2)!}e^{k\lambda} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{k\lambda} \end{bmatrix}\begin{bmatrix} x_2' \\ x_3' \\ \vdots \\ x_{m+1}' \end{bmatrix}$$

$$= \mathbf{C}\begin{bmatrix} e^{k\lambda} & \frac{k}{1!}e^{k\lambda} & \cdots & \frac{k^m}{m!}e^{k\lambda} \\ 0 & e^{k\lambda} & \cdots & \frac{k^{m-1}}{(m-1)!}e^{k\lambda} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{k\lambda} \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ \vdots \\ x_{m+1} + \frac{c_1'}{c_1}x_{m+1}' \end{bmatrix}$$

$$+ \begin{bmatrix} c_1 & c_2 & \cdots & c_m \end{bmatrix}\begin{bmatrix} e^{k\lambda} & \frac{k}{1!}e^{k\lambda} & \cdots & \frac{k^{m-1}}{(m-1)!}e^{k\lambda} \\ 0 & e^{k\lambda} & \cdots & \frac{k^{m-2}}{(m-2)!}e^{k\lambda} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{k\lambda} \end{bmatrix}\begin{bmatrix} x_1 + \frac{c_1'}{c_1}x_1' \\ x_2 + \frac{c_1'}{c_1}x_2' \\ \vdots \\ x_m + \frac{c_1'}{c_1}x_m' \end{bmatrix}$$

$$+ \begin{bmatrix} c_2' - \frac{c_1'}{c_1}c_2 & c_3' - \frac{c_1'}{c_1}c_3 & \cdots & c_{m+1}' - \frac{c_1'}{c_1}c_{m+1} \end{bmatrix}\begin{bmatrix} e^{k\lambda} & \frac{k}{1!}e^{k\lambda} & \frac{k}{2!}e^{k\lambda} & \cdots & \frac{k^{m-1}}{(m-1)!}e^{k\lambda} \\ 0 & e^{k\lambda} & \frac{k}{1!}e^{k\lambda} & \cdots & \frac{k^{m-2}}{(m-2)!}e^{k\lambda} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & e^{k\lambda} \end{bmatrix}\begin{bmatrix} x_2' \\ x_3' \\ \vdots \\ x_{m+1}' \end{bmatrix}$$

$$= \mathbf{C}\begin{bmatrix} e^{k\lambda} & \frac{k}{1!}e^{k\lambda} & \cdots & \frac{k^m}{m!}e^{k\lambda} \\ 0 & e^{k\lambda} & \cdots & \frac{k^{m-1}}{(m-1)!}e^{k\lambda} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{k\lambda} \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ \vdots \\ x_{m+1} + \frac{c_1'}{c_1}x_{m+1}' \end{bmatrix}$$

$$+ \begin{bmatrix} c_1 & c_2 & \cdots & c_m \end{bmatrix}\begin{bmatrix} e^{k\lambda} & \frac{k}{1!}e^{k\lambda} & \cdots & \frac{k^{m-1}}{(m-1)!}e^{k\lambda} \\ 0 & e^{k\lambda} & \cdots & \frac{k^{m-2}}{(m-2)!}e^{k\lambda} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{k\lambda} \end{bmatrix}\left(\begin{bmatrix} x_1 + \frac{c_1'}{c_1}x_1' \\ x_2 + \frac{c_1'}{c_1}x_2' \\ \vdots \\ x_m + \frac{c_1'}{c_1}x_m' \end{bmatrix} + \begin{bmatrix} t_{1,1}' & t_{1,2}' & \cdots & t_{1,m}' \\ 0 & t_{2,2}' & \cdots & t_{2,m}' \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & t_{m,m}' \end{bmatrix}\begin{bmatrix} x_2' \\ x_3' \\ \vdots \\ x_{m+1}' \end{bmatrix}\right) \tag{7.33}$$

$$= \mathbf{C}\begin{bmatrix} e^{k\lambda} & \frac{k}{1!}e^{k\lambda} & \cdots & \frac{k^m}{m!}e^{k\lambda} \\ 0 & e^{k\lambda} & \cdots & \frac{k^{m-1}}{(m-1)!}e^{k\lambda} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{k\lambda} \end{bmatrix}\left(\begin{bmatrix} 0 \\ 0 \\ \vdots \\ x_{m+1} + \frac{c_1'}{c_1}x_{m+1}' \end{bmatrix} + \begin{bmatrix} x_1 + \frac{c_1'}{c_1}x_1' \\ x_2 + \frac{c_1'}{c_1}x_2' \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} t_{1,1}' & t_{1,2}' & \cdots & t_{1,m}' \\ 0 & t_{2,2}' & \cdots & t_{2,m}' \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}\begin{bmatrix} x_2' \\ x_3' \\ \vdots \\ x_{m+1}' \end{bmatrix}\right)$$

$$= \mathbf{C}\begin{bmatrix} e^{k\lambda} & \frac{k}{1!}e^{k\lambda} & \cdots & \frac{k^m}{m!}e^{k\lambda} \\ 0 & e^{k\lambda} & \cdots & \frac{k^{m-1}}{(m-1)!}e^{k\lambda} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{k}{1!}e^{k\lambda} \\ 0 & 0 & \cdots & e^{k\lambda} \end{bmatrix}\left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{m+1} \end{bmatrix} + \begin{bmatrix} \frac{c_1'}{c_1} & t_{1,1}' & \cdots & t_{1,m}' \\ 0 & \frac{c_1'}{c_1} & \cdots & t_{2,m}' \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{c_1'}{c_1} \end{bmatrix}\begin{bmatrix} x_1' \\ x_2' \\ \vdots \\ x_{m+1}' \end{bmatrix}\right)$$

where (7.33) follows from the induction hypothesis. The lemma is true. □

Now, we are ready to prove Lemma 2.2.

Proof of Lemma 2.2. The proof is an induction on m, the size of matrix $\mathbf{A_c}$. Recall that here $\mathbf{C}$ can be a general matrix, so we use the definitions of $\mathbf{A_c}, \mathbf{C}$ given as (2.28), (2.29).

(i) When $m = 1$,

In this case, the system is scalar, and the lemma is trivially true. A rigorous proof goes as follows: Since $(\mathbf{A_c}, \mathbf{C})$ is observable, we can find a $1 \times l$ matrix $\mathbf{L}$ such that $\mathbf{LC}$ is not zero. Then, $(\mathbf{A_c}, \mathbf{LC})$ is observable, and the lemma is reduced to Lemma 7.13.

(ii) We will assume that the lemma holds for $(m-1)$-dimensional systems as an induction hypothesis, and prove the lemma holds for m.

The proof goes in three steps. First, we reduce the system to a system with scalar observations to apply Lemma 7.13. Then, we estimate one of the states, and subtract the estimation from the system — this procedure is known as successive decoding in information theory. Now, the system reduces to an $(m-1)$-dimensional one, so we apply the induction hypothesis.

To do this, we define $\mathbf{x} := \begin{bmatrix} \mathbf{x_{1,1}} \\ \mathbf{x_{1,2}} \\ \vdots \\ \mathbf{x}_{\mu,\nu_\mu} \end{bmatrix}$ where $\mathbf{x_{i,j}}$ are $m_{i,j} \times 1$ vectors, and $(\mathbf{x}_{\mathbf{1},\nu_\mathbf{1}})_{m_{1,\nu_1}}$ as the m_{1,ν_1}th element of $\mathbf{x}_{\mathbf{1},\nu_\mathbf{1}}$. We also define $(\mathbf{x})_k$ as the kth element of a vector $\mathbf{x}$ in general. Here, $\mathbf{x}$ can be thought as the states of the system. We first decode $(\mathbf{x}_{\mathbf{1},\nu_\mathbf{1}})_{m_{1,\nu_1}}$, and decode the remaining elements in $\mathbf{x}$.

• Reduction to Systems with Scalar Observations: By Lemma 7.13, we already know that the lemma is true for systems with scalar observations. Therefore, we will reduce a general system with vector observations to a system with scalar observations.

Claim 7.2. *There exist* $\mathbf{L}, \mathbf{C}', \mathbf{A}', \mathbf{x}'$ *that satisfy the following conditions.*
(i) $\mathbf{L}$ *is a* $1 \times l$ *row vector.*
(ii) $\mathbf{A}'$ *is a* $m' \times m'$ *square matrix given in Jordan form. The eigenvalues of* $\mathbf{A}'$ *belong to* $\{\lambda_1 + j\omega_1, \cdots, \lambda_\mu + j\omega_\mu\}$ *which is the set of eigenvalues of* $\mathbf{A}$*. The first Jordan block of* $\mathbf{A}'$ *is equal to* $\mathbf{A}_{\mathbf{1},\nu_\mathbf{1}}$*.*
(iii) $\mathbf{C}'$ *is a* $l \times m'$ *matrix and* $(\mathbf{A}', \mathbf{LC}')$ *is observable.*
(iv) $\mathbf{x}'$ *is a* $m' \times l$ *column vector.* $(\mathbf{x}')_{m_{1,\nu_1}} = (\mathbf{x}_{\mathbf{1},\nu_\mathbf{1}})_{m_{1,\nu_1}}$*.*
(v) $\mathbf{LC}e^{-k\mathbf{A_c}}\mathbf{x} = \mathbf{LC}'e^{-k\mathbf{A}'}\mathbf{x}'$*.*

What this claim implies is the following. By multiplying the matrix $\mathbf{L}$ to the vector observations, we can obtain scalar observations. However, the resulting system may not be observable any more. Therefore, we will carefully design the $\mathbf{L}$ matrix and reduced system matrices $\mathbf{A}'$, $\mathbf{C}'$, so that the system remains observable even with a scalar observation and the information about $(\mathbf{x}_{\mathbf{1},\nu_\mathbf{1}})_{m_{1,\nu_1}}$ remains intact. Furthermore, since the reduced system $(\mathbf{A}', \mathbf{LC}')$ has a scalar observation, all eigenvalues of $\mathbf{A}'$ has to be distinct to make the reduced system observable.

Proof. Since the first columns of $\mathbf{C_{1,1}}, \mathbf{C_{1,2}}, \cdots, \mathbf{C}_{\mathbf{1},\nu_\mathbf{1}}$ are linearly independent, there exists a $1 \times l$

matrix $\mathbf{L}$ such that the first elements of $\mathbf{LC_{1,1}}, \mathbf{LC_{1,2}}, \cdots, \mathbf{LC_{1,\nu_1-1}}$ are zeros and the first element of $\mathbf{LC_{1,\nu_1}}$ is non-zero. Then, we can observe that

$$\mathbf{LC}e^{-k\mathbf{A_c}}\mathbf{x} = \mathbf{L}\begin{bmatrix}\mathbf{C_{1,1}} & \cdots & \mathbf{C}_{\mu,\nu_\mu}\end{bmatrix}\begin{bmatrix} e^{-k\mathbf{A_{1,1}}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{-k\mathbf{A}_{\mu,\nu_\mu}} \end{bmatrix}\begin{bmatrix}\mathbf{x_{1,1}} \\ \vdots \\ \mathbf{x}_{\mu,\nu_\mu}\end{bmatrix}$$

$$= \mathbf{LC_{1,1}}e^{-k\mathbf{A_{1,1}}}\mathbf{x_{1,1}} + \mathbf{LC_{1,2}}e^{-k\mathbf{A_{1,2}}}\mathbf{x_{1,2}} + \cdots + \mathbf{LC}_{\mu,\nu_\mu}e^{-k\mathbf{A}_{\mu,\nu_\mu}}\mathbf{x}_{\mu,\nu_\mu} \quad (7.34)$$

Recall that the Jordan blocks $\mathbf{A_{i,1}}, \cdots, \mathbf{A_{i,\nu_i}}$ correspond to the same eigenvalue. We will merge these Jordan blocks into one Jordan block. However, since the size of Jordan blocks $\mathbf{A_{i,1}}, \cdots, \mathbf{A_{i,\nu_i}}$ are distinct, we will extend the small Jordan blocks to the size of the largest one by adding zero elements. Let the dimension of $\mathbf{A_{i,\bar{\nu}_i}}$ be the largest among $\mathbf{A_{i,1}}, \cdots, \mathbf{A_{i,\nu_i}}$, and $m_{i,\bar{\nu}_i}$ be the corresponding dimension. Then, we define $\mathbf{\bar{C}_{i,j}}$ as a matrix where the first $m_{i,\bar{\nu}_i} - m_{i,j}$ vectors are all zeros, and the remaining vectors are the same as those of $\mathbf{C_{i,j}}$. $\mathbf{\bar{A}_{i,j}}$ is defined as the same matrix as $\mathbf{A_{i,\bar{\nu}_i}}$. $\mathbf{\bar{x}_{i,j}}$ is defined as a column vector whose first $m_{i,\bar{\nu}_i} - m_{i,j}$ elements are all zeros, and the remaining elements are those of $\mathbf{x_{i,j}}$.

Then, by the construction, we know

$$(7.34) = \mathbf{L\bar{C}_{1,1}}e^{-k\mathbf{\bar{A}_{1,1}}}\mathbf{\bar{x}_{1,1}} + \mathbf{L\bar{C}_{1,2}}e^{-k\mathbf{\bar{A}_{1,2}}}\mathbf{\bar{x}_{1,2}} + \cdots + \mathbf{L\bar{C}}_{\mu,\nu_\mu}e^{-k\mathbf{\bar{A}}_{\mu,\nu_\mu}}\mathbf{\bar{x}}_{\mu,\nu_\mu}.$$

Furthermore, $\mathbf{A_{1,\nu_1}} = \mathbf{\bar{A}_{1,\nu_1}}$, $\mathbf{C_{1,\nu_1}} = \mathbf{\bar{C}_{1,\nu_1}}$, $\mathbf{x_{1,\nu_1}} = \mathbf{\bar{x}_{1,\nu_1}}$. The first elements of $\mathbf{LC_{1,1}}$, $\mathbf{LC_{1,2}}$, $\cdots$, $\mathbf{LC_{1,\nu_1-1}}$ are zeros and the first element of $\mathbf{LC_{1,\nu_1}}$ is non-zero.

Now, we get the same dimension systems $(\mathbf{\bar{A}_{i,1}}, \mathbf{L\bar{C}_{i,1}}), \cdots, (\mathbf{\bar{A}_{i,\nu_i}}, \mathbf{L\bar{C}_{i,\nu_i}})$. However, none of them might be observable. Thus, we will truncate the matrices to make sure that at least one of them is observable. Recall that since $\mathbf{L\bar{C}_{i,j}}$ is a row vector and $\mathbf{\bar{A}_{i,j}}$ is a single Jordan block, the system is observable as long as the first element of $\mathbf{L\bar{C}_{i,j}}$ is not zero. Thus, we will truncate the matrices until we see at least one nonzero element among the first elements of $\mathbf{L\bar{C}_{i,1}}, \cdots, \mathbf{L\bar{C}_{i,\nu_i}}$. Let k_i be the smallest number such that at least one of the k_ith elements of $\mathbf{L\bar{C}_{i,1}}, \cdots, \mathbf{L\bar{C}_{i,\nu_i}}$ becomes nonzero, and let $\mathbf{L\bar{C}_{i,\nu_i^\star}}$ be the vector that achieves the minimum.

Then, we will reduce the dimensions of $(\mathbf{\bar{A}_{i,j}}, \mathbf{L\bar{C}_{i,j}})$ by truncating the first $(k_i - 1)$ vectors. Define $\mathbf{C'_{i,j}}$ as the matrix obtained by removing the first $(k_i - 1)$ columns from $\mathbf{\bar{C}_{i,j}}$, $\mathbf{A'_{i,j}}$ as a vector obtained by removing the first $(k_i - 1)$ rows and columns from $\mathbf{\bar{A}_{i,j}}$, and $\mathbf{x'_{i,j}}$ as a vector obtained by removing the first $(k_i - 1)$ elements from $\mathbf{\bar{x}_{i,j}}$.

Then, by construction, the resulting systems $(\mathbf{A'_{i,\nu_i^\star}}, \mathbf{LC'_{i,\nu_i^\star}})$ are observable. We can also see that $\nu_1^\star = \nu_1$, $\mathbf{C'_{1,\nu_1^\star}} = \mathbf{\bar{C}_{1,\nu_1}} = \mathbf{C_{1,\nu_1}}$, $\mathbf{A'_{1,\nu_1^\star}} = \mathbf{\bar{A}_{1,\nu_1}} = \mathbf{A_{1,\nu_1}}$, and $\mathbf{x'_{1,\nu_1^\star}} = \mathbf{\bar{x}_{1,\nu_1}} = \mathbf{x_{1,\nu_1}}$. In words, the Jordan block $\mathbf{A_{1,\nu_1}}$ was not affected by the above manipulations. Moreover, by construction, the first elements of $\mathbf{LC'_{1,1}}, \cdots, \mathbf{LC'_{1,\nu_1-1}}$ are all zero.

Denote $\mathbf{C'} := \begin{bmatrix}\mathbf{C'_{1,\nu_1^\star}} & \mathbf{C'_{2,\nu_2^\star}} & \cdots & \mathbf{C'}_{\mu,\nu_\mu^\star}\end{bmatrix}$ and $\mathbf{A'} := diag\{\mathbf{A'_{1,\nu_1^\star}}, \mathbf{A'_{2,\nu_2^\star}}, \cdots, \mathbf{A'}_{\mu,\nu_\mu^\star}\}$. Then, (7.34) can be written as follows:

$$
\begin{aligned}
(7.34) &= \mathbf{LC}'_{\mathbf{1},\mathbf{1}} e^{-k\mathbf{A}'_{1,1}} \mathbf{x}'_{\mathbf{1},\mathbf{1}} + \mathbf{LC}'_{\mathbf{1},\mathbf{2}} e^{-k\mathbf{A}'_{1,2}} \mathbf{x}'_{\mathbf{1},\mathbf{2}} + \cdots + \mathbf{LC}'_{\mu,\nu_\mu} e^{-k\mathbf{A}'_{\mu,\nu_\mu}} \mathbf{x}'_{\mu,\nu_\mu} \\
&= \mathbf{LC}'_{\mathbf{1},\nu^\star_{\mathbf{1}}} e^{-k\mathbf{A}'_{1,\nu^\star_1}} (\mathbf{x}'_{\mathbf{1},\nu^\star_{\mathbf{1}}} + \sum_{j\in\{1,\cdots,\nu_1\}\setminus\nu^\star_1} \mathbf{T}_{\mathbf{1},\mathbf{j}} \mathbf{x}'_{\mathbf{1},\mathbf{j}}) + \cdots \\
&\quad + \mathbf{LC}'_{\mu,\nu^\star_\mu} e^{-k\mathbf{A}'_{\mu,\nu^\star_\mu}} (\mathbf{x}'_{\mu,\nu^\star_\mu} + \sum_{j\in\{1,\cdots,\nu_\mu\}\setminus\nu^\star_\mu} \mathbf{T}_{\mu,\mathbf{j}} \mathbf{x}'_{\mu,\mathbf{j}}) \qquad (7.35) \\
&= \begin{bmatrix} \mathbf{LC}'_{\mathbf{1},\nu^\star_{\mathbf{1}}} & \cdots & \mathbf{LC}'_{\mathbf{1},\nu^\star_\mu} \end{bmatrix}
\begin{bmatrix} e^{-k\mathbf{A}'_{1,\nu^\star_1}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{-k\mathbf{A}'_{\mu,\nu^\star_\mu}} \end{bmatrix}
\underbrace{\begin{bmatrix} \mathbf{x}'_{\mathbf{1},\nu^\star_{\mathbf{1}}} + \sum_{j\in\{1,\cdots,\nu_1\}\setminus\nu^\star_1} \mathbf{T}_{\mathbf{1},\mathbf{j}} \mathbf{x}'_{\mathbf{1},\mathbf{j}} \\ \vdots \\ \mathbf{x}'_{\mu,\nu^\star_\mu} + \sum_{j\in\{1,\cdots,\nu_\mu\}\setminus\nu^\star_\mu} \mathbf{T}_{\mu,\mathbf{j}} \mathbf{x}'_{\mu,\mathbf{j}} \end{bmatrix}}_{:=\mathbf{x}'} \\
&= \mathbf{LC}' e^{-k\mathbf{A}'} \mathbf{x}'
\end{aligned}
$$

where (7.35) follows from Lemma 7.14. Here, we can easily see that $\mathbf{A}'$ satisfies the condition (ii) of the claim, and $(\mathbf{A}', \mathbf{LC}')$ is observable since each $(\mathbf{A}'_{\mathbf{i},\nu^\star_{\mathbf{i}}}, \mathbf{LC}'_{\mathbf{i},\nu^\star_{\mathbf{i}}})$ is observable.

Moreover, by Lemma 7.14, we know that $\mathbf{T}_{\mathbf{1},\mathbf{1}}, \cdots, \mathbf{T}_{\mathbf{1},\nu_{\mathbf{1}}-\mathbf{1}}$ are upper triangular matrices whose diagonal elements are zeros. Therefore, $(\mathbf{x}')_{m_{1,\nu_1}} = (\mathbf{x}'_{\mathbf{1},\nu_{\mathbf{1}}})_{m_{1,\nu_1}} = (\mathbf{x}_{\mathbf{1},\nu_{\mathbf{1}}})_{m_{1,\nu_1}}$. Therefore, the condition (iv) of the claim is also satisfied. □

• Decoding $(\mathbf{x}_{\mathbf{1},\nu_{\mathbf{1}}})_{m_{1,\nu_1}}$: Now, we reduced the system to a system with a scalar observation. Then, we can apply Lemma 7.13 to decode $(\mathbf{x}_{\mathbf{1},\nu_{\mathbf{1}}})_{m_{1,\nu_1}}$.

Claim 7.3. *We can find a polynomial $p'(k)$ and a family of stopping times $\{S'(\epsilon,k) : k \in \mathbb{Z}^+, \epsilon > 0\}$ such that for all $\epsilon > 0$, $k \in \mathbb{Z}^+$ there exist $k \le k_1 < k_2 < \cdots < k_{m'} \le S'(\epsilon,k)$ and $\mathbf{M}'_{\mathbf{1}}$ satisfying:*
(i) $\beta[k_i] = 1$ for $1 \le i \le m'$
(ii) $\mathbf{M}'_{\mathbf{1}} \begin{bmatrix} \mathbf{L} & 0 & \cdots & 0 \\ 0 & \mathbf{L} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{L} \end{bmatrix} \begin{bmatrix} \mathbf{C}e^{-(k_1 I + t_{k_1})\mathbf{A}} \\ \mathbf{C}e^{-(k_2 I + t_{k_2})\mathbf{A}} \\ \vdots \\ \mathbf{C}e^{-(k_{m'} I + t_{k_{m'}})\mathbf{A}} \end{bmatrix} \mathbf{x} = (\mathbf{x}_{\mathbf{1},\nu_{\mathbf{1}}})_{m_{1,\nu_1}}$
(iii) $|\mathbf{M}'_{\mathbf{1}}|_{max} \le \frac{p'(S'(\epsilon,k))}{\epsilon} e^{\lambda_1 S'(\epsilon,k) I}$
(iv) $\lim_{\epsilon\downarrow 0} \exp \limsup_{s\to\infty} \sup_{k\in\mathbb{Z}^+} \frac{1}{s} \log \mathbb{P}\{S'(\epsilon,k) - k = s\} \le p_e$.

This claim is showing that there exists an estimator $\mathbf{M}'_{\mathbf{1}} diag\{\mathbf{L}, \cdots, \mathbf{L}\}$ which can estimate the state $(\mathbf{x}_{\mathbf{1},\nu_{\mathbf{1}}})_{m_{1,\nu_1}}$ with observations at time $k_1, \cdots, k_m$.

Proof. By construction, $(\mathbf{A}', \mathbf{LC}')$ is observable and $\mathbf{LC}'$ is a row vector. Thus, by Lemma 7.13 we can find a polynomial $p'(k)$ and a family of stopping times $\{S'(\epsilon,k) : k \in \mathbb{Z}^+, \epsilon > 0\}$ such that for all $\epsilon > 0$, $k \in \mathbb{Z}^+$ there exist $k \le k_1 < k_2 < \cdots < k_{m'} \le S'(\epsilon,k)$ and $\mathbf{M}'$ satisfying:
(i) $\beta[k_i] = 1$ for $1 \le i \le m'$

(ii) $\mathbf{M}' \begin{bmatrix} \mathbf{LC}' e^{-(k_1 I + t_{k_1})\mathbf{A}'} \\ \mathbf{LC}' e^{-(k_2 I + t_{k_2})\mathbf{A}'} \\ \vdots \\ \mathbf{LC}' e^{-(k_{m'} I + t_{k_{m'}})\mathbf{A}'} \end{bmatrix} = \mathbf{I}$

(iii) $|\mathbf{M}'|_{max} \leq \frac{p'(S'(\epsilon,k))}{\epsilon} e^{\lambda_1 S'(\epsilon,k) I}$

(iv) $\lim_{\epsilon \downarrow 0} \exp \limsup_{s \to \infty} \sup_{k \in \mathbb{Z}^+} \frac{1}{s} \log \mathbb{P}\{S'(\epsilon,k) - k = s\} \leq p_e$.

Let $\mathbf{M}'_1$ be the m_{1,ν_1}th row of $\mathbf{M}'$. Then,

$$\mathbf{M}'_1 \begin{bmatrix} \mathbf{L} & 0 & \cdots & 0 \\ 0 & \mathbf{L} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{L} \end{bmatrix} \begin{bmatrix} \mathbf{C} e^{-(k_1 I + t_{k_1})\mathbf{A_c}} \\ \mathbf{C} e^{-(k_2 I + t_{k_2})\mathbf{A_c}} \\ \vdots \\ \mathbf{C} e^{-(k_{m'} I + t k_{m'})\mathbf{A_c}} \end{bmatrix} \mathbf{x} = \mathbf{M}'_1 \begin{bmatrix} \mathbf{LC} e^{-(k_1 I + t_{k_1})\mathbf{A_c}} \mathbf{x} \\ \mathbf{LC} e^{-(k_2 I + t_{k_2})\mathbf{A_c}} \mathbf{x} \\ \vdots \\ \mathbf{LC} e^{-(k_{m'} I + t_{k_{m'}})\mathbf{A_c}} \mathbf{x} \end{bmatrix}$$

$$= \mathbf{M}'_1 \begin{bmatrix} \mathbf{L}'\mathbf{C}' e^{-(k_1 I + t_{k_1})\mathbf{A}'} \mathbf{x}' \\ \mathbf{L}'\mathbf{C}' e^{-(k_2 I + t_{k_2})\mathbf{A}'} \mathbf{x}' \\ \vdots \\ \mathbf{L}'\mathbf{C}' e^{-(k_{m'} I + t_{k_{m'}})\mathbf{A}'} \mathbf{x}' \end{bmatrix} (\because \textit{Claim } 7.2 \text{ (v)})$$

$$= \mathbf{M}'_1 \begin{bmatrix} \mathbf{L}'\mathbf{C}' e^{-(k_1 I + t_{k_1})\mathbf{A}'} \\ \mathbf{L}'\mathbf{C}' e^{-(k_2 I + t_{k_2})\mathbf{A}'} \\ \vdots \\ \mathbf{L}'\mathbf{C}' e^{-(k_{m'} I + t_{k_{m'}})\mathbf{A}'} \end{bmatrix} \mathbf{x}' = (\mathbf{x}')_{m_{1,\nu_1}} = (\mathbf{x}_{\mathbf{1},\nu_\mathbf{1}})_{m_{1,\nu_1}} (\because \textit{Claim } 7.2 \text{ (iv)}).$$

□

• Subtracting $(\mathbf{x}_{\mathbf{1},\nu_\mathbf{1}})_{m_{1,\nu_1}}$ from the observations: Now, we have an estimate for $(\mathbf{x}_{\mathbf{1},\nu_\mathbf{1}})_{m_{1,\nu_1}}$. We will remove it from the system. $\mathbf{A}''$,$\mathbf{C}''$ and $\mathbf{x}''$ are the system matrices after the removal. Formally, $\mathbf{A}''$,$\mathbf{C}''$ and $\mathbf{x}''$ are obtained by removing $\sum_{1 \leq i \leq \nu_i} m_{1,i}$th row and column from $\mathbf{A_c}$, removing $\sum_{1 \leq i \leq \nu_i} m_{1,i}$th row from $\mathbf{C}$ and removing $\sum_{1 \leq i \leq \nu_i} m_{1,i}$th component from $\mathbf{x}$ respectively.

Obviously, $\mathbf{A}'' \in \mathbb{C}^{(m-1)\times(m-1)}$ and $\mathbf{C}'' \in \mathbb{C}^{l \times (m-1)}$. Moreover, since the last element of the Jordan block $\mathbf{A}_{\mathbf{1},\nu_\mathbf{1}}$ is removed and the observability only depends on the first element, $(\mathbf{A}'', \mathbf{C}'')$ is observable. Denote $\lambda''_1 + \omega''_1$ be the eigenvalue of $\mathbf{A}''$ with the largest real part. Then, trivially $\lambda''_1 \leq \lambda_1$.

The new system $(\mathbf{A}'', \mathbf{C}'')$ and the original system $(\mathbf{A}, \mathbf{C})$ are related as follows. Denote the $\sum_{1 \leq i \leq \nu_i} m_{1,i}$th column of $\mathbf{C} e^{-k\mathbf{A_c}}$ as $\mathbf{R}(k)$. Then, we have

$$\mathbf{C} \mathbf{e}^{-\mathbf{k}\mathbf{A_c}} \mathbf{x} - \mathbf{R}(k)(\mathbf{x}_{\mathbf{1},\nu_\mathbf{1}})_{m_{1,\nu_1}} = \mathbf{C}'' \mathbf{e}^{-\mathbf{k}\mathbf{A}''} \mathbf{x}'' \tag{7.36}$$

which can be easily proved from the block diagonal structure of $\mathbf{A_c}$. We can further see that there exists a polynomial $p'''(k)$ such that $|\mathbf{R}(k)|_{max} \leq p'''(k) e^{-k\lambda_1}$.

• Decoding the remaining element of $\mathbf{x}$: We decoded and subtracted the state $(\mathbf{x_{1,\nu_1}})_{m_{1,\nu_1}}$ from the system. Now, we can apply the induction hypothesis to the remaining $(m-1)$-dimensional system and estimate the remaining states.

By the induction hypothesis, for given $S'(\epsilon, k)$, we can find $m'' \in \mathbb{Z}$ and a polynomial $p''(k)$ and a family of stopping time $\{S''(\epsilon, S'(\epsilon, k)) : S'(\epsilon, k) \in \mathbb{Z}^+, 0 < \epsilon < 1\}$ such that for all $0 < \epsilon < 1$ there exist $S'(\epsilon, k) < k_{m'+1} < \cdots < k_{m''} \leq S''(\epsilon, S'(\epsilon, k))$ and a $(m-1) \times (m''-m')l$ matrix $\mathbf{M}''$ satisfying the following conditions:

(i) $\beta[k_i] = 1$ for $m'+1 \leq i \leq m''$

(ii) $\mathbf{M}'' \begin{bmatrix} \mathbf{C}''e^{-(k_{m'+1}I + t_{k_{m'+1}})\mathbf{A}''} \\ \mathbf{C}''e^{-(k_{m'+2}I + t_{k_{m'+2}})\mathbf{A}''} \\ \vdots \\ \mathbf{C}''e^{-(k_{m''}I + t_{k_{m''}})\mathbf{A}''} \end{bmatrix} = \mathbf{I}_{(m-1)\times(m-1)}$

(iii) $|\mathbf{M}''|_{max} \leq \frac{p''(S''(\epsilon, S'(\epsilon,k)))}{\epsilon} e^{\lambda_1'' S''(\epsilon, S'(\epsilon,k))I}$

(iv) $\lim_{\epsilon \downarrow 0} \exp \limsup_{s\to\infty} \operatorname{ess\,sup} \frac{1}{s} \log \mathbb{P}\{S''(\epsilon, S'(\epsilon, k)) - S'(\epsilon, k) = s | \mathcal{F}_{S'(\epsilon,k)}\} \leq p_e$

where $\mathcal{F}_n$ is the σ-field generated by $\beta[0], \cdots, \beta[n]$ and $t_0, \cdots, t_n$.

Then,

$$\mathbf{x}'' = \mathbf{M}'' \begin{bmatrix} \mathbf{C}''e^{-(k_{m'+1}I+t_{k_{m'+1}})\mathbf{A}''} \\ \mathbf{C}''e^{-(k_{m'+2}I+t_{k_{m'+2}})\mathbf{A}''} \\ \vdots \\ \mathbf{C}''e^{-(k_{m''}I+t_{k_{m''}})\mathbf{A}''} \end{bmatrix} \mathbf{x}''$$

$$= \mathbf{M}'' \begin{bmatrix} \mathbf{C}e^{-(k_{m'+1}I+t_{k_{m'+1}})\mathbf{A_c}}\mathbf{x} - \mathbf{R}(k_{m'+1}I + t_{k_{m'+1}})(\mathbf{x_{1,\nu_1}})_{m_{1,\nu_1}} \\ \mathbf{C}e^{-(k_{m'+2}I+t_{k_{m'+2}})\mathbf{A_c}}\mathbf{x} - \mathbf{R}(k_{m'+2}I + t_{k_{m'+2}})(\mathbf{x_{1,\nu_1}})_{m_{1,\nu_1}} \\ \vdots \\ \mathbf{C}e^{-(k_{m''}I+t_{k_{m''}})\mathbf{A_c}}\mathbf{x} - \mathbf{R}(k_{m''}I + t_{k_{m''}})(\mathbf{x_{1,\nu_1}})_{m_{1,\nu_1}} \end{bmatrix} \tag{7.37}$$

$$= \mathbf{M}'' \left(\begin{bmatrix} \mathbf{C}e^{-(k_{m'+1}I+t_{k_{m'+1}})\mathbf{A_c}} \\ \mathbf{C}e^{-(k_{m'+2}I+t_{k_{m'+2}})\mathbf{A_c}} \\ \vdots \\ \mathbf{C}e^{-(k_{m''}I+t_{k_{m''}})\mathbf{A_c}} \end{bmatrix} \mathbf{x} - \begin{bmatrix} \mathbf{R}(k_{m'+1}I + t_{k_{m'+1}}) \\ \mathbf{R}(k_{m'+2}I + t_{k_{m'+2}}) \\ \vdots \\ \mathbf{R}(k_{m''}I + t_{k_{m''}}) \end{bmatrix} (\mathbf{x_{1,\nu_1}})_{m_{1,\nu_1}} \right)$$

$$= \mathbf{M}'' \left(\begin{bmatrix} \mathbf{C}e^{-(k_{m'+1}I+t_{k_{m'+1}})\mathbf{A_c}} \\ \mathbf{C}e^{-(k_{m'+2}I+t_{k_{m'+2}})\mathbf{A_c}} \\ \vdots \\ \mathbf{C}e^{-(k_{m''}I+t_{k_{m''}})\mathbf{A_c}} \end{bmatrix} \mathbf{x} - \begin{bmatrix} \mathbf{R}(k_{m'+1}I + t_{k_{m'+1}}) \\ \mathbf{R}(k_{m'+2}I + t_{k_{m'+2}}) \\ \vdots \\ \mathbf{R}(k_{m''}I + t_{k_{m''}}) \end{bmatrix} \mathbf{M_1'} \begin{bmatrix} \mathbf{L} & 0 & \cdots & 0 \\ 0 & \mathbf{L} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{L} \end{bmatrix} \begin{bmatrix} \mathbf{C}e^{-(k_1I+t_{k_1})\mathbf{A_c}} \\ \mathbf{C}e^{-(k_2I+t_{k_2})\mathbf{A_c}} \\ \vdots \\ \mathbf{C}e^{-(k_{m'}I+t_{k_{m'}})\mathbf{A_c}} \end{bmatrix} \mathbf{x} \right) \tag{7.38}$$

$$= \mathbf{M}'' \left[- \begin{bmatrix} \mathbf{R}(k_{m'+1}I + t_{k_{m'+1}}) \\ \mathbf{R}(k_{m'+2}I + t_{k_{m'+2}}) \\ \vdots \\ \mathbf{R}(k_{m''}I + t_{k_{m''}}) \end{bmatrix} \mathbf{M_1'} \begin{bmatrix} \mathbf{L} & 0 & \cdots & 0 \\ 0 & \mathbf{L} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{L} \end{bmatrix} \quad \mathbf{I} \right] \begin{bmatrix} \mathbf{C}e^{-(k_1I+t_{k_1})\mathbf{A_c}} \\ \mathbf{C}e^{-(k_2I+t_{k_2})\mathbf{A_c}} \\ \vdots \\ \mathbf{C}e^{-(k_{m''}I+t_{k_{m''}})\mathbf{A_c}} \end{bmatrix} \mathbf{x}$$

where (7.37) follows from (7.36), and (7.38) follows from the condition (ii) of Claim 7.3. Therefore, we can recover the remaining states of $\mathbf{x}$.

Moreover, we have

$$\left|\mathbf{M}'' \begin{bmatrix} -\begin{bmatrix} \mathbf{R}(k_{m'+1}I + t_{k_{m'+1}}) \\ \mathbf{R}(k_{m'+2}I + t_{k_{m'+2}}) \\ \vdots \\ \mathbf{R}(k_{m''}I + t_{k_{m''}}) \end{bmatrix} \mathbf{M_1'} \begin{bmatrix} \mathbf{L} & 0 & \cdots & 0 \\ 0 & \mathbf{L} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{L} \end{bmatrix} & \mathbf{I} \end{bmatrix}\right|_{max}$$

$$\lesssim |\mathbf{M}''|_{max} \cdot \max\left\{ \left|\begin{bmatrix} \mathbf{R}(k_{m'+1}I + t_{k_{m'+1}}) \\ \mathbf{R}(k_{m'+2}I + t_{k_{m'+2}}) \\ \vdots \\ \mathbf{R}(k_{m''}I + t_{k_{m''}}) \end{bmatrix}\right|_{max} |\mathbf{M_1'}|_{max} \, |\mathbf{L}|_{max}, 1 \right\}$$

$$\lesssim \frac{p''(S''(\epsilon, S'(\epsilon,k)))}{\epsilon} e^{\lambda_1'' S''(\epsilon, S'(\epsilon,k))I}$$

$$\cdot \max\left\{ p'''(k_{m''}I + t_{k_{m''}}) e^{-\lambda_1(k_{m'+1}I + t_{k_{m'+1}})} \cdot \frac{p'(S'(\epsilon,k))}{\epsilon} e^{\lambda_1 S'(\epsilon,k)I} \cdot |\mathbf{L}|_{max}, 1 \right\}$$

$$\lesssim \frac{\bar{p}(S''(\epsilon, S'(\epsilon,k)))}{\epsilon^2} e^{\lambda_1 S''(\epsilon, S'(\epsilon,k))I} \;\; (\because S'(\epsilon,k) < k_{m'+1} < k_{m''} \le S''(\epsilon, S'(\epsilon,k)),\; \lambda_1'' \le \lambda_1)$$

for some polynomial $\bar{p}(k)$. Since for some $\bar{\bar{p}}(k)$

$$|\mathbf{M_1'}|_{max} \le \frac{p'(S'(\epsilon,k))}{\epsilon} e^{\lambda_1 S'(\epsilon,k)I} \le \frac{\bar{\bar{p}}(S''(\epsilon, S'(\epsilon,k)))}{\epsilon^2} e^{\lambda_1 S''(\epsilon, S'(\epsilon,k))I}$$

and we can recover $\mathbf{x}$ from $\mathbf{x}''$ and $(\mathbf{x_{1,\nu_1}})_{m_{1,\nu_1}}$, there exists $\mathbf{M}$ and a polynomial $p(k)$ such that

$$\mathbf{M} \begin{bmatrix} \mathbf{C}e^{-(k_1 I + t_{k_1})\mathbf{A_c}} \\ \vdots \\ \mathbf{C}e^{-(k_{m''} I + t_{k_{m''}})\mathbf{A_c}} \end{bmatrix} = \mathbf{I}_{m\times m}$$

and

$$|\mathbf{M}|_{max} \le \frac{p(S''(\epsilon, S'(\epsilon,k)))}{\epsilon^2} e^{\lambda_1 S''(\epsilon, S'(\epsilon,k))I}.$$

Moreover, since

$$\lim_{\epsilon\downarrow 0} \exp \limsup_{s\to\infty} \operatorname{ess\,sup} \log \mathbb{P}\{S''(\epsilon, S'(\epsilon,k)) - S'(\epsilon,k) | \mathcal{F}_{S'(\epsilon,k)}\} \le p_e$$

and

$$\lim_{\epsilon\downarrow 0} \exp \limsup_{s\to\infty} \sup_{k\in\mathbb{Z}^+} \log \mathbb{P}\{S'(\epsilon,k) - k\} \le p_e,$$

by Lemma 7.2

$$\lim_{\epsilon\downarrow 0} \exp \limsup_{s\to\infty} \sup_{k\in\mathbb{Z}^+} \log \mathbb{P}\{S''(\epsilon, S'(\epsilon,k)) - k\} \le p_e.$$

Therefore, by putting $S(\epsilon,k) := S''(\epsilon^{\frac{1}{2}}, S'(\epsilon^{\frac{1}{2}}, k))$, $S(\epsilon,k)$ satisfies all the conditions of the lemma.

□

7.5 Lemmas about the Observability Gramian of Discrete-Time Systems

Now, we will consider the discrete-time systems discussed in Section 2.6. Like the continuous time case, we start from a simpler case when $\mathbf{C}$ is a row vector and $\mathbf{A}$ has no eigenvalue cycles. The definitions corresponding to (2.35) for the row vector case are given as follows: Let $\mathbf{A}$ be a $m \times m$ Jordan form matrix and $\mathbf{C}$ be $1 \times m$ row vector which can be written as

$$\mathbf{A} = diag\{\mathbf{A_{1,1}}, \mathbf{A_{1,2}}, \cdots, \mathbf{A_{1,\nu_1}}, \cdots, \mathbf{A_{\mu,1}}, \cdots, \mathbf{A_{\mu,\nu_\mu}}\} \tag{7.39}$$

$$\mathbf{C} = \begin{bmatrix}\mathbf{C_{1,1}}, \mathbf{C_{1,2}}, \cdots, \mathbf{C_{1,\nu_1}}, \cdots, \mathbf{C_{\mu,1}}, \cdots, \mathbf{C_{\mu,\nu_\mu}}\end{bmatrix} \tag{7.40}$$

where

$\mathbf{A_{i,j}}$ is a Jordan block with eigenvalue $\lambda_{i,j} e^{j2\pi\omega_{i,j}}$ and size $m_{i,j}$

$m_{i,1} \leq m_{i,2} \leq \cdots \leq m_{i,\nu_i}$ for all $i = 1, \cdots, \mu$

$\lambda_{i,1} = \lambda_{i,2} = \cdots = \lambda_{i,\nu_i}$ for all $i = 1, \cdots, \mu$

$\lambda_{1,1} > \lambda_{2,1} > \cdots > \lambda_{\mu,1} \geq 1$

$\{\lambda_{i,1}, \cdots, \lambda_{i,\nu_i}\}$ is cycle with length ν_i and period p_i

For all $(i,j) \neq (i',j')$, $\omega_{i,j} - \omega_{i',j'} \notin \mathbb{Q}$

$\mathbf{C_{i,j}}$ is a $1 \times m_{i,j}$ complex matrix and its first element is non-zero

$\lambda_i e^{j2\pi\omega_i}$ is (i,i) element of $\mathbf{A}$.

Here, we can notice that $\mathbf{A}$ has no eigenvalue cycles since $\omega_{i,j} - \omega_{i',j'} \notin \mathbb{Q}$ for all $(i,j) \neq (i',j')$, and $\mathbf{C}$ is a row vector. By Theorem 2.6, the condition that the first elements of $\mathbf{C_{i,j}}$ are non-zero corresponds to the observability condition of $(\mathbf{A}, \mathbf{C})$ since $\mathbf{C}$ is a row vector.

Let's state lemmas which parallel Lemma 7.6 and Lemma 7.8. In fact, the proofs of the lemmas are very similar to those of Lemma 7.6 and Lemma 7.8 and we omit the proofs here.

Lemma 7.15. *Let* $\mathbf{A}$ *and* $\mathbf{C}$ *be given as* (7.39) *and* (7.40). *Then, there exists a polynomial* $p(k)$ *such that for all* $\epsilon > 0$ *and* $0 \leq k_1 \leq \cdots \leq k_m$, *if*

$$\left| \det\left(\begin{bmatrix} \mathbf{CA}^{-k_1} \\ \mathbf{CA}^{-k_2} \\ \vdots \\ \mathbf{CA}^{-k_m} \end{bmatrix} \right) \right| \geq \epsilon \prod_{1 \leq i \leq m} \lambda_i^{-k_i}$$

then

$$\left\| \begin{bmatrix} \mathbf{CA}^{-k_1} \\ \mathbf{CA}^{-k_2} \\ \vdots \\ \mathbf{CA}^{-k_m} \end{bmatrix} \right\|_{max} \leq \frac{p(k_m)}{\epsilon} \lambda_1^{k_m}$$

Proof. It can be easily proved in a similar way to Lemma 7.6. □

Lemma 7.16. *Let* $\mathbf{A}$ *and* $\mathbf{C}$ *be given as* (7.39) *and* (7.40)*. Define* $a_{i,j}$ *and* $C_{i,j}$ *as the* (i,j) *element and cofactor of* $\begin{bmatrix} \mathbf{CA}^{-k_1} \\ \mathbf{CA}^{-k_2} \\ \vdots \\ \mathbf{CA}^{-k_m} \end{bmatrix}$ *respectively. Then there exists* $g_\epsilon(k) : \mathbb{R}^+ \to \mathbb{R}^+$ *and* $a \in \mathbb{R}^+$ *such that for all* $\epsilon > 0$ *and* $k_1, \cdots, k_m$ *satisfying*

$$(i) 0 \leq k_1 < k_2 < \cdots < k_m$$
$$(ii) k_m - k_{m-1} \geq g_\epsilon(k_{m-1})$$
$$(iii) g_\epsilon(k) \leq a(1 + \log(k+1))$$
$$(iv) \left| \sum_{m-m_\mu+1 \leq i \leq m} a_{m,i} C_{m,i} \right| \geq \epsilon \prod_{1 \leq i \leq m} {\lambda_i}^{-k_i}$$

the following inequality holds:

$$\left| \det \left(\begin{bmatrix} \mathbf{CA}^{-k_1} \\ \mathbf{CA}^{-k_2} \\ \vdots \\ \mathbf{CA}^{-k_m} \end{bmatrix} \right) \right| \geq \frac{1}{2} \epsilon \prod_{1 \leq i \leq m} {\lambda_i}^{-k_i}.$$

Proof. It can be easily proved in a similar way to Lemma 7.8. □

Like the continuous-time case, these lemmas reduce questions about the inverse of the observability Gramian to questions about the determinant of the observability Gramian.

7.6 Uniform Convergence of Sequences satisfying Weyl's criterion (Discrete-Time Systems)

As we did in the continuous-time case, we will prove that the determinant of the observability matrix is large enough regardless of the erasure pattern. The main difference from the continuous-time case of Appendix 7.3 is the measure that must be used. While we used the Lebesgue measure to measure the bad event —the event that the determinant of the observability matrix is small—, we use the counting measure in this section.

The main idea of this section is approximating aperiodic deterministic sequences by random variables using ergodic theory [54]. The necessary and sufficient condition for a sequence to behave like uniformly distributed random variables in $[0,1]$ is known as Weyl's criterion. We first state a general ergodic theorem, and derive the Weyl's criterion as a corollary.

Theorem 7.2 (Koksma and Szusz inequality [54])**.** *Consider a s-dimensional sequence* $\mathbf{x_1}, \mathbf{x_2}, \cdots \in \mathbb{R}^s$*, and let* $\alpha := (\alpha_1, \cdots, \alpha_s)$ *and* $\beta := (\beta_1, \cdots, \beta_s)$*. For any positive integer* m*, we have*

$$\sup_{0 \le \alpha_i < \beta_i \le 1} \left| \frac{A([\alpha, \beta); N, \{\mathbf{x_n}\})}{N} - \prod_{1 \le i \le s} (\beta_i - \alpha_i) \right|$$
$$\le 2s^2 3^{s+1} \left(\frac{1}{m} + \sum_{\mathbf{h} \in \mathbb{Z}^s, 0 < |\mathbf{h}|_\infty \le m} \frac{1}{r(\mathbf{h})} \left| \frac{1}{N} \sum_{1 \le n \le N} e^{2\pi\sqrt{-1}\langle \mathbf{h}, \mathbf{x_n} \rangle} \right| \right)$$

where

$$A([\alpha, \beta); N, \{\mathbf{x_n}\}) := \sum_{1 \le n \le N} \mathbf{1}\{\mathbf{x_n} \in [\alpha_1, \beta_1) \times [\alpha_2, \beta_2) \cdots \times [\alpha_s, \beta_s)\} \tag{7.41}$$
$$r(\mathbf{h}) := \prod_{1 \le j \le s} \max\{|h_j|, 1\}.$$

Proof. See [54] for the proof. ☐

Here, we can see $A([\alpha, \beta); N, \{\mathbf{x_n}\})$ is the counting measure of the event that a sequence falls in the set $[\alpha, \beta)$. The theorem tells us that the counting measure is close to the Lebesgue measure of the set $[\alpha, \beta)$ uniformly over all α, β.

Using this theorem, we can easily derive[2] the Weyl's criterion for a family of sequences.

Definition 7.2. *Consider a family of s-dimensional sequences* $\mathcal{J} = \{(\mathbf{x_{1,\sigma}}, \mathbf{x_{2,\sigma}}, \cdots) : \sigma \in J, x_{i,\sigma} \in \mathbb{R}^s\}$*. Here, the index set for the sequences,* J*, can be infinite. If for all* $\mathbf{h} \in \mathbb{Z}^s \setminus \{\mathbf{0}\}$*,*

$$\lim_{N \to \infty} \sup_{\sigma \in \mathcal{J}} \left| \frac{1}{N} \sum_{1 \le n \le N} e^{j2\pi\langle \mathbf{h}, \mathbf{x_{n,\sigma}} \rangle} \right| = 0$$

then the family of sequences is said to satisfy Weyl's criterion.

Theorem 7.3 (Weyl's criterion [54])**.** *Consider a family of s-dimensional sequences* $\mathcal{J} = \{(\mathbf{x_{1,\sigma}}, \mathbf{x_{2,\sigma}}, \cdots) : \sigma \in J, x_{i,\sigma} \in \mathbb{R}^s\}$*, which satisfy the Weyl's criterion. Then, this family of sequences satisfies*

$$\lim_{N \to \infty} \sup_{\sigma \in \mathcal{J}} \sup_{0 \le \alpha_i < \beta_i \le 1} \left| \frac{A([\alpha, \beta); N, \{\mathbf{x_{n,\sigma}}\})}{N} - \prod_{1 \le i \le s} (\beta_i - \alpha_i) \right| = 0,$$

[2] The original Weyl's criterion [54] is shown for only one sequence. Here we extend Weyl's criterion to a family of sequences. For this, we state a generalized theorem of the Weyl's criterion and prove it.

where the definition of $A([\alpha,\beta);N,\{\mathbf{x}_{\mathbf{n},\sigma}\})$ *is given in* (7.41).

Proof. By Theorem 7.2, for any positive integer m, we have

$$\begin{aligned}
&\sup_{\sigma\in\mathcal{J}}\sup_{0\le\alpha_i<\beta_i\le 1}\left|\frac{A([\alpha,\beta);N,\{\mathbf{x}_{\mathbf{n},\sigma}\})}{N}-\prod_{1\le i\le s}(\beta_i-\alpha_i)\right| \\
&\le \sup_{\sigma\in\mathcal{J}} 2s^2 3^{s+1}\left(\frac{1}{m}+\sum_{0<|\mathbf{h}|_\infty\le m}\frac{1}{r(\mathbf{h})}\left|\frac{1}{N}\sum_{1\le n\le N} e^{2\pi j\langle \mathbf{h},\mathbf{x}_{\mathbf{n},\sigma}\rangle}\right|\right) \qquad (7.42)
\end{aligned}$$

To prove the theorem, it is enough to show that for all $\delta>0$ there exists N' such that for all $N>N'$

$$\sup_{\sigma\in\mathcal{J}}\sup_{0\le\alpha_i<\beta_i\le 1}\left|\frac{A([\alpha,\beta);N,\{\mathbf{x}_{\mathbf{n},\sigma}\})}{N}-\prod_{1\le i\le s}(\beta_i-\alpha_i)\right|<\delta. \qquad (7.43)$$

Let's choose $m:=\frac{4s^2 3^{s+1}}{\delta}$ so that

$$\frac{2s^2 3^{s+1}}{m}<\frac{\delta}{2}. \qquad (7.44)$$

Once we fix m, there are only $(2m+1)^s$ number of $\mathbf{h}\in\mathbb{Z}^s$ such that $|\mathbf{h}|_\infty\le m$. Furthermore, by the definition of Weyl's criterion, we can find N'' such that for all $N>N''$,

$$\sup_{\sigma\in\mathcal{J}}\left|\frac{1}{N}\sum_{1\le n\le N} e^{j2\pi\langle \mathbf{h},\mathbf{x}_{\mathbf{n},\sigma}\rangle}\right|<\frac{1}{(2m+1)^s 2s^2 3^{s+1}}\frac{\delta}{2}.$$

Thus, we can find N'' such that for all $N>N''$ the following holds:

$$2s^2 3^{s+1} s^{m+1}\max_{0<|\mathbf{h}|_\infty\le m}\sup_{\sigma\in\mathcal{J}}\left|\frac{1}{N}\sum_{1\le n\le N} e^{j2\pi\langle \mathbf{h},\mathbf{x}_{\mathbf{n},\sigma}\rangle}\right|<\frac{\delta}{2}. \qquad (7.45)$$

Therefore, by plugging (7.44), (7.45) into (7.42), we can prove (7.43). Thus, the theorem is true. □

Since we are mainly interested in the fractional part of sequences, it will be helpful to denote $\langle x\rangle := x-\lfloor x\rfloor$. Although $\langle \mathbf{x},\mathbf{y}\rangle$ is the inner product between two vectors, these two definitions can be distinguished by counting the number of arguments. Let's consider some specific sequences, and see whether they satisfy Weyl's criterion.

Example 7.1. $(\langle\sqrt{2}n\rangle,\langle\sqrt{3}n\rangle)$ *satisfies Weyl's criterion and* $(\langle\sqrt{2}n\rangle,\langle(\sqrt{2}+\sqrt{3})n\rangle)$ *does too.* $(\langle\sqrt{2}n\rangle,\langle(\sqrt{2}+0.5)n\rangle)$ *does not satisfy Weyl's criterion and neither does* $\left(\langle\sqrt{2}n\rangle,\left\langle\frac{\sqrt{2}}{2}n\right\rangle\right)$.

Therefore, among general sequences in the form of $(\langle\omega_1 n\rangle,\langle\omega_2 n\rangle,\cdots,\langle\omega_m n\rangle)$, there are sequences which satisfy Weyl's criterion and others do not. However, the following lemma reveals all sequences can be written as linear combinations of basis sequences which satisfy Weyl's criterion. This idea is very similar to that linear-algebraic concepts like linear decomposition and basis.

Lemma 7.17. *Consider an m-dimensional sequence $(\langle\omega_1 n\rangle, \langle\omega_2 n\rangle, \cdots, \langle\omega_m n\rangle)$. Then, there exists $k \le m$ and $p \in \mathbb{N}$ such that*

$$\omega_i = \frac{q_{i,0}}{p} + \sum_{1\le j\le k} q_{i,j}\gamma_j$$

where

$$q_{i,j} \in \mathbb{Z},$$
$$(\langle\gamma_1 n\rangle, \langle\gamma_2 n\rangle, \cdots, \langle\gamma_k n\rangle) \text{ satisfies Weyl's criterion.}$$

Proof. Before the proof, we can observe the following two facts.

First, since as long as $\langle \mathbf{h}, \mathbf{w}\rangle$ is not an integer,

$$\frac{1}{N}\sum_{1\le n\le N} e^{j2\pi\langle \mathbf{h},(\langle\omega_1 n\rangle,\langle\omega_2 n\rangle,\cdots,\langle\omega_m n\rangle)\rangle} = \frac{1}{N}\frac{e^{j2\pi(h_1\omega_1+h_2\omega_2+\cdots+h_m\omega_m)}\left(1-e^{j2\pi N(h_1\omega_1+h_2\omega_2+\cdots+h_m\omega_m)}\right)}{1-e^{j2\pi(h_1\omega_1+h_2\omega_2+\cdots+h_m\omega_m)}},$$

the statement that the sequence $(\langle\omega_1 n\rangle, \langle\omega_2 n\rangle, \cdots, \langle\omega_m n\rangle)$ does not satisfy Weyl's criterion is equivalent to there being $h_1, h_2, \cdots, h_m \in \mathbb{Z}$ that are not identically zero and make

$$h_1\omega_1 + h_2\omega_2 + \cdots + h_m\omega_m \in \mathbb{Z}. \tag{7.46}$$

The second observation is that if $(\langle\omega_1 n\rangle, \langle\omega_2 n\rangle, \cdots, \langle\omega_m n\rangle)$ satisfies Weyl's criterion then for all $a_1, \cdots, a_m \in \mathbb{N}$, $(\left\langle\frac{\omega_1}{a_1}n\right\rangle, \left\langle\frac{\omega_2}{a_2}n\right\rangle, \cdots, \left\langle\frac{\omega_m}{a_m}n\right\rangle)$ also satisfies Weyl's criterion. To see this, suppose $(\left\langle\frac{\omega_1}{a_1}n\right\rangle, \left\langle\frac{\omega_2}{a_2}n\right\rangle, \cdots, \left\langle\frac{\omega_m}{a_m}n\right\rangle)$ did not satisfy Weyl's criterion. Then, by (7.46) there would exist $(h_1, h_2, \cdots, h_m) \in \mathbb{Z}^m \setminus \{\mathbf{0}\}$ such that $h_1\frac{\omega_1}{a_1} + h_2\frac{\omega_2}{a_2} + \cdots + h_m\frac{\omega_m}{a_m} \in \mathbb{Z}$. So, $\frac{h_1\prod_{1\le i\le m}a_i}{a_1}\omega_1 + \frac{h_2\prod_{1\le i\le m}a_i}{a_2}\omega_2 + \cdots + \frac{h_m\prod_{1\le i\le m}a_i}{a_m}\omega_m \in \mathbb{Z}$ as well as $(\frac{h_1\prod_{1\le i\le m}a_i}{a_1}, \cdots, \frac{h_m\prod_{1\le i\le m}a_i}{a_m}) \in \mathbb{Z}^m \setminus \{\mathbf{0}\}$. But since $(\langle\omega_1 n\rangle, \langle\omega_2 n\rangle, \cdots, \langle\omega_m n\rangle)$ would not satisfy Weyl's criterion, this causes a contradiction.

Now, we will prove the lemma by induction on m.

(i) When $m = 1$,

If $\langle\omega_1 n\rangle$ satisfies Weyl's criterion, the lemma is trivially true by selecting $\gamma_1 = \omega_1$ and $q_{1,1} = 1$. If $\langle\omega_1 n\rangle$ does not satisfy Weyl's criterion, then by (7.46), ω_1 is a rational number. So we can find $q_{1,0}$ and p such that $\omega_1 = \frac{q_{1,0}}{p}$, and set the $k = 0$.

(ii) Assume that the lemma is true for $m - 1$.

If $(\langle\omega_1 n\rangle, \langle\omega_2 n\rangle, \cdots, \langle\omega_m n\rangle)$ satisfies Weyl's criterion, the lemma follows by selecting $k = m$, $\gamma_i = \omega_i$ and $q_{i,i} = 1$.

If $(\langle\omega_1 n\rangle, \langle\omega_2 n\rangle, \cdots, \langle\omega_m n\rangle)$ does not satisfy Weyl's criterion, by (7.46) there exists $(h_1, h_2, \cdots, h_m) \in \mathbb{Z}^m \setminus \{\mathbf{0}\}$ and $h \in \mathbb{Z}$ such that $h_1\omega_1 + h_2\omega_2 + \cdots + h_m\omega_m = h$. Without loss of generality, let's say $h_1 \ne 0$. Then

$$\omega_1 = -\frac{h_2}{h_1}\omega_2 - \frac{h_3}{h_1}\omega_3 - \cdots - \frac{h_m}{h_1}\omega_m + \frac{h}{h_1}. \tag{7.47}$$

By the induction hypothesis, we know that there exists $k' \leq m-1$, $p' \in \mathbb{N}$, $q'_{i,j} \in \mathbb{Z}$, γ'_i such that

$$\begin{aligned}\omega_2 &= \frac{q'_{2,0}}{p'} + \sum_{1 \leq j \leq k'} q'_{2,j}\gamma'_j \\ &\vdots \\ \omega_m &= \frac{q'_{m,0}}{p'} + \sum_{1 \leq j \leq k'} q'_{m,j}\gamma'_j. \end{aligned} \tag{7.48}$$

where $(\langle \gamma'_1 n \rangle, \langle \gamma'_2 n \rangle, \cdots, \langle \gamma'_k n \rangle)$ satisfies Weyl's criterion. Therefore, by plugging (7.48) to (7.47) we can find $q'_{1,j} \in \mathbb{Z}$ such that

$$\omega_1 = \frac{q'_{1,0}}{|h_1 \cdot p'|} + \sum_{1 \leq i \leq k} q'_{1,i}\frac{\gamma'_i}{h_1}.$$

By the second observation, $(\left\langle \frac{\gamma'_1}{h_1} n \right\rangle, \left\langle \frac{\gamma'_2}{h_1} n \right\rangle, \cdots, \left\langle \frac{\gamma'_k}{h_1} n \right\rangle)$ satisfies Weyl's criterion, so we can use $p = |h_1 \cdot p'|$ and $\gamma_i = \frac{\gamma'_i}{h_1}$ to show that the lemma also holds for m.

Therefore, by induction the lemma is true. □

Now, we can decompose the sequences into basis sequences which satisfy Weyl's criterion, and so behave like uniform random variables. The main difference from the uniform convergence discussion of Appendix 7.3 is the number of random variables. In other words, in continuous-time systems with random jitter, only one random variable is introduced at each sample for the random jitter. However, this is not the case in discrete-time systems.

Let $\mathbf{A_1} = \begin{bmatrix} e^{j\sqrt{2}} & 0 \\ 0 & e^{j2\sqrt{2}} \end{bmatrix}$, $\mathbf{A_2} = \begin{bmatrix} e^{j\sqrt{2}} & 0 \\ 0 & e^{j\sqrt{3}} \end{bmatrix}$, $\mathbf{C} = \begin{bmatrix} 1 & 1 \end{bmatrix}$. The row of the observability gramian of $(\mathbf{A_1}, \mathbf{C})$ is $\mathbf{CA_1}^n = \begin{bmatrix} e^{j\sqrt{2}n} & e^{j2\sqrt{2}n} \end{bmatrix}$. In this case, the elements of $\mathbf{CA_1}^n$ do not satisfy Weyl's criterion. Thus, it can be approximated by $\begin{bmatrix} e^{jX} & e^{j2X} \end{bmatrix}$ where X is uniform in $[0, 2\pi]$. This involves only one random variable.

However, the row of the observability gramian of $(\mathbf{A_2}, \mathbf{C})$ is $\mathbf{CA_2}^n = \begin{bmatrix} e^{j\sqrt{2}n} & e^{j\sqrt{3}n} \end{bmatrix}$ whose elements satisfy Weyl's criterion. Thus, it can be approximated by $\begin{bmatrix} e^{jX_1} & e^{jX_2} \end{bmatrix}$ where X_1, X_2 are independent uniform random variables in $[0, 2\pi]$, and so involves two random variables.

Therefore, the lemmas derived in Appendix 7.3 have to be generalized to multiple random variables, and then the multiple random variables can be used to model deterministic sequences.

Intuitively, adding more randomness should not cause any problems, so generalization to multiple random variables must be possible. We first extend Lemma 7.10 which was written for a single random variable to multiple random variables.

Lemma 7.18. *Let* $\mathbf{X}$ *be* $(X_1, X_2, \cdots, X_\nu)$ *where* X_i *are i.i.d. random variables whose distribution is uniform between* 0 *and* 2π*. Let* $\mathbf{k_1}, \mathbf{k_2}, \cdots, \mathbf{k_\mu} \in \mathbb{R}^\nu$ *be distinct. Then, for strictly positive* γ*,* Γ

$(\gamma \le \Gamma)$, *and* $m \in \{1, \cdots, \mu\}$

$$\sup_{|a_m| \ge \gamma, |a_i| \le \Gamma, a_i \in \mathbb{C}} \mathbb{P}\{|\sum_{i=1}^{\mu} a_i e^{j<\mathbf{k_i}, \mathbf{X}>}| < \epsilon\} \to 0 \ \textit{as}\ \epsilon \downarrow 0.$$

Proof. We will prove the lemma by induction on ν, the number of random variables.

(i) When $\nu = 1$. The lemma reduces to Lemma 7.10.

(ii) Let's assume the lemma is true for $1, \cdots, \nu - 1$.

Without loss of generality, we can assume $m = 1$ by symmetry. We will prove the lemma by dividing into cases based on $\mathbf{k_i}$. Let the jth component of $\mathbf{k_i}$ be denoted as k_{ij}.

First, consider the case when $k_{1,1} = k_{2,1} = \cdots = k_{\mu,1}$. Then,

$$\begin{aligned}
&\sup_{|a_1| \ge \gamma, |a_i| \le \Gamma} \mathbb{P}\{|\sum_{i=1}^{\mu} a_i e^{j<\mathbf{k_i}, \mathbf{X}>}| < \epsilon\} = \sup_{|a_1| \ge \gamma, |a_i| \le \Gamma} \mathbb{P}\{|\sum_{i=1}^{\mu} a_i e^{j \sum_{1 \le j \le \nu} k_{i,j} X_j}| < \epsilon\} \\
&= \sup_{|a_1| \ge \gamma, |a_i| \le \Gamma} \mathbb{P}\{|e^{j k_{1,1} X_1}| \cdot |\sum_{i=1}^{\mu} a_i e^{j \sum_{2 \le j \le \nu} k_{i,j} X_j}| < \epsilon\} \\
&= \sup_{|a_1| \ge \gamma, |a_i| \le \Gamma} \mathbb{P}\{|\sum_{i=1}^{\mu} a_i e^{j \sum_{2 \le j \le \nu} k_{i,j} X_j}| < \epsilon\} \to 0 \ (\because \text{induction hypothesis}).
\end{aligned}$$

Second, consider the case when $k_{i,1} \ne k_{j,1}$ for some i, j. Without loss of generality, we can assume that $k_{1,1} = k_{2,1} = \cdots = k_{\mu_1,1}$ and $k_{1,1} \ne k_{j,1}$ for all $\mu_1 < j \le \mu$. Then, for all $\epsilon' > 0$, we have

$$\begin{aligned}
&\sup_{|a_1| \ge \gamma, |a_i| \le \Gamma} \mathbb{P}\{|\sum_{i=1}^{\mu} a_i e^{j<\mathbf{k_i}, \mathbf{X}>}| < \epsilon\} \\
&= \sup_{|a_1| \ge \gamma, |a_i| \le \Gamma} \mathbb{P}\{|\sum_{i=1}^{\mu_1} a_i e^{j<\mathbf{k_i}, \mathbf{X}>} + \sum_{i=\mu_1+1}^{\mu} a_i e^{j<\mathbf{k_i}, \mathbf{X}>}| < \epsilon\} \\
&\le \sup_{|a_1| \ge \gamma, |a_i| \le \Gamma} \mathbb{P}\{|\sum_{i=1}^{\mu_1} a_i e^{j<\mathbf{k_i}, \mathbf{X}>} + \sum_{i=\mu_1+1}^{\mu} a_i e^{j<\mathbf{k_i}, \mathbf{X}>}| < \epsilon \Big| |\sum_{i=1}^{\mu_1} a_i e^{j \sum_{2 \le j \le \nu} k_{i,j} X_j}| \ge \epsilon'\} \\
&+ \mathbb{P}\{|\sum_{i=1}^{\mu_1} a_i e^{j \sum_{2 \le j \le \nu} k_{i,j} X_j}| < \epsilon'\} \\
&= \sup_{|a_1| \ge \gamma, |a_i| \le \Gamma} \mathbb{P}\{|(\sum_{i=1}^{\mu_1} a_i e^{j \sum_{2 \le j \le \nu} k_{i,j} X_j}) e^{j k_{1,1} X_1} + \sum_{i=\mu_1+1}^{\mu} a_i e^{j<\mathbf{k_i}, \mathbf{X}>}| < \epsilon \Big| |\sum_{i=1}^{\mu_1} a_i e^{j \sum_{2 \le j \le \nu} k_{i,j} X_j}| \ge \epsilon'\} \\
&+ \mathbb{P}\{|\sum_{i=1}^{\mu_1} a_i e^{j \sum_{2 \le j \le \nu} k_{i,j} X_j}| < \epsilon'\} \\
&\le \sup_{|a'_1| \ge \epsilon', |a'_i| \le \mu\Gamma} \mathbb{P}_{X_1}\{|a'_1 e^{j k_{1,1} X_1} + \sum_{i=\mu_1+1}^{\mu} a'_i e^{j k_{i,1} X_1}| < \epsilon\} + \sup_{|a_1| \ge \gamma, |a_i| \le \Gamma} \mathbb{P}\{|\sum_{i=1}^{\mu_1} a_i e^{j \sum_{2 \le j \le \nu} k_{i,j} X_j}| < \epsilon'\}.
\end{aligned}$$

Therefore, by the induction hypothesis (since the first term has only one random variable, and the second term has $\nu - 1$ random variables)

$$\lim_{\epsilon \to 0} \sup_{|a_1| \geq \gamma, |a_i| \leq \Gamma} \mathbb{P}\{|\sum_{i=1}^{\mu} a_i e^{j<\mathbf{k_i}, \mathbf{X}>}| < \epsilon\}$$

$$\leq \lim_{\epsilon' \to 0} \lim_{\epsilon \to 0} \sup_{|a_1'| \geq \epsilon', |a_i'| \leq \mu\Gamma} \mathbb{P}\{|a_1' e^{jk_{1,1}X_1} + \sum_{i=\mu_1+1}^{\mu} a_i' e^{jk_{i,1}X_1}| < \epsilon\} + \sup_{|a_1| \geq \gamma, |a_i| \leq \Gamma} \mathbb{P}\{|\sum_{i=1}^{\mu_1} a_i e^{j\sum_{2 \leq j \leq \nu} k_{i,j} X_j}| < \epsilon'\}$$

$$= 0.$$

Therefore, the lemma is true. □

Now, we will consider a deterministic sequence in the form of $(< \omega_1 n >, \cdots, < \omega_\mu n >)$. As we have shown in Lemma 7.17, this sequence can be thought of as a linear combination of basis sequences which satisfy Weyl's criterion. Thus, we can approximate the deterministic sequence as a linear combination of multiple uniform random variables considered in Lemma 7.18.

Lemma 7.19. *Let $\omega_1, \omega_2, \cdots, \omega_\mu$ be real numbers such that $\omega_i - \omega_j \notin \mathbb{Q}$ for all $i \neq j$. Then, for strictly positive numbers γ and Γ $(\gamma \leq \Gamma)$, and $m \in \{1, \cdots, \mu\}$*

$$\lim_{\epsilon \downarrow 0} \lim_{N \to \infty} \sup_{|a_m| \geq \gamma, |a_i| \leq \Gamma, k \in \mathbb{Z}} \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}\{|\sum_{i=1}^{\mu} a_i e^{j2\pi\omega_i(n+k)}| < \epsilon\} \to 0.$$

Proof. By Lemma 7.17, ω_i can be written as $\langle \mathbf{q_i}, \rho \rangle$ where $\mathbf{q_i} = (q_{i,0}, q_{i,1}, \cdots, q_{i,r}) \in \mathbb{Z}^{r+1}$, $\rho = (\frac{1}{s}, \rho_1, \cdots, \rho_r) \in \mathbb{R}^{r+1}$ and $s \in \mathbb{N}$. Here, $(\langle \rho_1 n \rangle, \langle \rho_2 n \rangle, \cdots, \langle \rho_r n \rangle)$ satisfies Weyl's criterion. Since $\omega_i - \omega_j \notin \mathbb{Q}$ for all $i \neq j$, $(q_{i,1}, q_{i,2}, \cdots, q_{i,r}) \neq (q_{j,1}, q_{j,2}, \cdots, q_{j,r})$.

For given $k, N, M \in \mathbb{N}$, and $m_1, \cdots, m_r \in \{1, \cdots, M\}$, define a set $S_{m_1,\cdots,m_r}$ as[3]

$$\left\{n \in \{1, \cdots, N\} : \frac{m_1 - 1}{M} \leq \langle \rho_1(n+k) \rangle < \frac{m_1}{M}, \cdots, \frac{m_r - 1}{M} \leq \langle \rho_r(n+k) \rangle < \frac{m_r}{M}\right\}.$$

Then, for all $k, N, M \in \mathbb{N}$ and $\epsilon > 0$, we have the following:

$$\sum_{n=1}^{N} \mathbf{1}\{|\sum_{i=1}^{\mu} a_i e^{j2\pi\omega_i(n+k)}| < \epsilon\}$$

$$= \sum_{n=1}^{N} \sum_{1 \leq m_1 \leq M, \cdots, 1 \leq m_r \leq M} \mathbf{1}\{|\sum_{i=1}^{\mu} a_i e^{j2\pi\omega_i(n+k)}| < \epsilon, n \in S_{m_1,\cdots,m_r}\}$$

$$\leq \sum_{n=1}^{N} \sum_{1 \leq m_1 \leq M, \cdots, 1 \leq m_r \leq M} \mathbf{1}\{\min_{n \in S_{m_1,\cdots,m_r}} |\sum_{i=1}^{\mu} a_i e^{j2\pi\omega_i(n+k)}| < \epsilon, n \in S_{m_1,\cdots,m_r}\}$$

$$= \sum_{n=1}^{N} \sum_{1 \leq m_1 \leq M, \cdots, 1 \leq m_r \leq M} \mathbf{1}\{\min_{n \in S_{m_1,\cdots,m_r}} |\sum_{i=1}^{\mu} a_i e^{j2\pi\omega_i(n+k)}| < \epsilon\} \cdot \mathbf{1}\{n \in S_{m_1,\cdots,m_r}\}. \quad (7.49)$$

[3]Notice that the definition of $S_{m_1,\cdots,m_r}$ also depends on k, N, M as well as $m_1, \cdots, m_r$. However, we omit the dependence on k, N, M in the definition for simplicity.

Moreover, we also know by the definitions of $\mathbf{q_i}$ and ρ,

$$\begin{aligned}\sum_{i=1}^{\mu} a_i e^{j2\pi\omega_i(n+k)} &= \sum_{i=1}^{\mu} a_i e^{j2\pi\langle \mathbf{q_i},\rho\rangle(n+k)} \\ &= \sum_{i=1}^{\mu} a_i e^{j2\pi\left(\frac{q_{i,0}}{s}(n+k)+q_{i,1}\rho_1(n+k)+\cdots+q_{i,r}\rho_r(n+k)\right)} \\ &= \sum_{i=1}^{\mu} a_i e^{j2\pi\left(\frac{q_{i,0}}{s}(n+k)+q_{i,1}\langle\rho_1(n+k)\rangle+\cdots+q_{i,r}\langle\rho_r(n+k)\rangle\right)} (\because q_{i,j}\in\mathbb{Z}).\end{aligned}$$

Thus, by defining $\mathbf{X_{m_1,\cdots,m_r}}$ as a random vector which is uniformly distributed over $[\frac{m_1-1}{M},\frac{m_1}{M})\times\cdots\times[\frac{m_r-1}{M},\frac{m_r}{M})$ and $\mathbf{q_i'}=(q_{i,1},q_{i,2},\cdots,q_{i,r})$, $\rho'=(\rho_1,\rho_2,\cdots,\rho_r)$, we can conclude

$$\begin{aligned}\max_{n\in S_{m_1,\cdots,m_r}} \Big|\sum_{i=1}^{\mu} a_i e^{j2\pi\omega_i(n+k)}\Big| &= \max_{n\in S_{m_1,\cdots,m_r}} \Big|\sum_{i=1}^{\mu} a_i e^{j2\pi\left(\frac{q_{i,0}}{s}(n+k)+q_{i,1}\langle\rho_1(n+k)\rangle+\cdots+q_{i,r}\langle\rho_r(n+k)\rangle\right)}\Big| \\ &\geq \Big|\sum_{i=1}^{\mu} a_i e^{j2\pi\left(\frac{q_{i,0}}{s}(n+k)+\left\langle\mathbf{q_i'},\mathbf{X_{m_1,\cdots,m_r}}\right\rangle\right)}\Big| \quad a.e. \end{aligned} \tag{7.50}$$

By (7.50), (7.49) can be upper bounded as follows:

$$\begin{aligned}(7.49) &\leq \sum_{n=1}^{N}\sum_{1\leq m_1\leq M,\cdots,1\leq m_r\leq M} \mathbb{P}\Big\{\min_{n\in S_{m_1,\cdots,m_r}}\Big|\sum_{i=1}^{\mu} a_i e^{j2\pi\omega_i(n+k)}\Big| - \max_{n\in S_{m_1,\cdots,m_r}}\Big|\sum_{i=1}^{\mu} a_i e^{j2\pi\omega_i(n+k)}\Big| \\ &+ \Big|\sum_{i=1}^{\mu} a_i e^{j2\pi\left(\frac{q_{i,0}}{s}(n+k)+\left\langle\mathbf{q_i'},\mathbf{X_{m_1,\cdots,m_r}}\right\rangle\right)}\Big| < \epsilon\Big\}\cdot\mathbf{1}\{n\in S_{m_1,\cdots,m_r}\} \\ &= \sum_{n=1}^{N}\sum_{1\leq m_1\leq M,\cdots,1\leq m_r\leq M} \mathbb{P}\Big\{\min_{n\in S_{m_1,\cdots,m_r}}\Big|\sum_{i=1}^{\mu} a_i e^{j2\pi\left(\frac{q_{i,0}}{s}(n+k)+\left\langle\mathbf{q_i'},\rho'\right\rangle(n+k)\right)}\Big| \\ &- \max_{n\in S_{m_1,\cdots,m_r}}\Big|\sum_{i=1}^{\mu} a_i e^{j2\pi\left(\frac{q_{i,0}}{s}(n+k)+\left\langle\mathbf{q_i'},\rho'\right\rangle(n+k)\right)}\Big| \\ &+ \Big|\sum_{i=1}^{\mu} a_i e^{j2\pi\left(\frac{q_{i,0}}{s}(n+k)+\left\langle\mathbf{q_i'},\mathbf{X_{m_1,\cdots,m_r}}\right\rangle\right)}\Big| < \epsilon\Big\}\cdot\mathbf{1}\{n\in S_{m_1,\cdots,m_r}\} \\ &\leq \sum_{n=1}^{N}\sum_{1\leq m_1\leq M,\cdots,1\leq m_r\leq M}\max_{0\leq s'<s} \mathbb{P}\Big\{\min_{n\in S_{m_1,\cdots,m_r}}\Big|\sum_{i=1}^{\mu} a_i e^{j2\pi\left(\frac{s'}{s}+\left\langle\mathbf{q_i'},\rho'\right\rangle(n+k)\right)}\Big| \\ &- \max_{n\in S_{m_1,\cdots,m_r}}\Big|\sum_{i=1}^{\mu} a_i e^{j2\pi\left(\frac{s'}{s}+\left\langle\mathbf{q_i'},\rho'\right\rangle(n+k)\right)}\Big| + \Big|\sum_{i=1}^{\mu} a_i e^{j2\pi\left(\frac{s'}{s}+\left\langle\mathbf{q_i'},\mathbf{X_{m_1,\cdots,m_r}}\right\rangle\right)}\Big| < \epsilon\Big\}\cdot\mathbf{1}\{n\in S_{m_1,\cdots,m_r}\} \\ &\leq \sum_{n=1}^{N}\sum_{1\leq m_1\leq M,\cdots,1\leq m_r\leq M}\sum_{0\leq s'<s} \mathbb{P}\Big\{\min_{n\in S_{m_1,\cdots,m_r}}\Big|\sum_{i=1}^{\mu} a_i e^{j2\pi\left(\frac{s'}{s}+\left\langle\mathbf{q_i'},\rho'\right\rangle(n+k)\right)}\Big| \\ &- \max_{n\in S_{m_1,\cdots,m_r}}\Big|\sum_{i=1}^{\mu} a_i e^{j2\pi\left(\frac{s'}{s}+\left\langle\mathbf{q_i'},\rho'\right\rangle(n+k)\right)}\Big| + \Big|\sum_{i=1}^{\mu} a_i e^{j2\pi\left(\frac{s'}{s}+\left\langle\mathbf{q_i'},\mathbf{X_{m_1,\cdots,m_r}}\right\rangle\right)}\Big| < \epsilon\Big\}\cdot\mathbf{1}\{n\in S_{m_1,\cdots,m_r}\}.\end{aligned} \tag{7.51}$$

Here, we have

$$\max_{n\in S_{m_1,\cdots,m_r}} |\sum_{i=1}^{\mu} a_i e^{j2\pi\left(\frac{s'}{s}+\left\langle \mathbf{q}'_\mathbf{i},\rho'\right\rangle(n+k)\right)}|$$

$$= \max_{n\in S_{m_1,\cdots,m_r}} |\sum_{i=1}^{\mu} a_i e^{j2\pi\left(\frac{s'}{s}+q_{i,1}\langle\rho_1(n+k)\rangle+\cdots+q_{i,r}\langle\rho_r(n+k)\rangle\right)}|(\because q_{i,j}\in\mathbb{Z})$$

$$\le \sup_{0\le\Delta_i<\frac{1}{M}} |\sum_{i=1}^{\mu} a_i e^{j2\pi\left(\frac{s'}{s}+q_{i,1}\frac{m_1-1}{M}+\cdots+q_{i,r}\frac{m_r-1}{M}+q_{i,1}\Delta_1+\cdots+q_{i,r}\Delta_r\right)}|$$

$$= \sup_{0\le\Delta_i<\frac{1}{M}} |\sum_{i=1}^{\mu} a_i e^{j2\pi\left(\frac{s'}{s}+q_{i,1}\frac{m_1-1}{M}+\cdots+q_{i,r}\frac{m_r-1}{M}\right)} + a_i e^{j2\pi\left(\frac{s'}{s}+q_{i,1}\frac{m_1-1}{M}+\cdots+q_{i,r}\frac{m_r-1}{M}\right)}$$

$$(-1+\cos 2\pi(q_{i,1}\Delta_1+\cdots+q_{i,r}\Delta_r)+j\sin 2\pi(q_{i,1}\Delta_1+\cdots+q_{i,r}\Delta_r))|$$

$$\le |\sum_{i=1}^{\mu} a_i e^{j2\pi\left(\frac{s'}{s}+q_{i,1}\frac{m_1-1}{M}+\cdots+q_{i,r}\frac{m_r-1}{M}\right)}|$$

$$+\sum_{i=1}^{\mu} |a_i e^{j2\pi\left(\frac{s'}{s}+q_{i,1}\frac{m_1-1}{M}+\cdots+q_{i,r}\frac{m_r-1}{M}\right)}|$$

$$\cdot(\sup_{0\le\Delta_i<\frac{1}{M}} |-1+\cos 2\pi(q_{i,1}\Delta_1+\cdots+q_{i,r}\Delta_r)| + \sup_{0\le\Delta_i<\frac{1}{M}} |\sin 2\pi(q_{i,1}\Delta_1+\cdots+q_{i,r}\Delta_r)|)$$

$$\le |\sum_{i=1}^{\mu} a_i e^{j2\pi\left(\frac{s'}{s}+q_{i,1}\frac{m_1-1}{M}+\cdots+q_{i,r}\frac{m_r-1}{M}\right)}| + 4\pi\sum_{i=1}^{\mu}|a_i| \sup_{0\le\Delta_i<\frac{1}{M}} |q_{i,1}\Delta_1+\cdots+q_{i,r}\Delta_r| \tag{7.52}$$

$$\le |\sum_{i=1}^{\mu} a_i e^{j2\pi\left(\frac{s'}{s}+q_{i,1}\frac{m_1-1}{M}+\cdots+q_{i,r}\frac{m_r-1}{M}\right)}| + \frac{4\pi\Gamma}{M}\sum_{i=1}^{\mu}\sum_{j=1}^{r}|q_{i,j}|.(\because \text{We assumed } |a_i|\le\Gamma)$$

where (7.52) comes from the fact that $|\sin x|\le|x|$ and $|-1+\cos x|\le|x|$ for all $x\in\mathbb{R}$.

Likewise, we also have

$$\min_{n\in S_{m_1,\cdots,m_r}} |\sum_{i=1}^{\mu} a_i e^{j2\pi\left(\frac{s'_i}{s}+\left\langle \mathbf{q}'_\mathbf{i},\rho'\right\rangle(n+k)\right)}|$$

$$\ge |\sum_{i=1}^{\mu} a_i e^{j2\pi\left(\frac{s'_i}{s}+q_{i,1}\frac{m_1-1}{M}+\cdots+q_{i,r}\frac{m_r-1}{M}\right)}| - \frac{4\pi\Gamma}{M}\sum_{i=1}^{\mu}\sum_{j=1}^{r}|q_{i,j}|.$$

Therefore,

$$\sup_{\frac{m_i-1}{M}\le\langle\rho_i(n+k)\rangle<\frac{m_i}{M}} |\sum_{i=1}^{\mu} a_i e^{j2\pi\left(\frac{s'_i}{s}+\left\langle \mathbf{q}'_\mathbf{i},\rho'\right\rangle(n+k)\right)}| - \inf_{\frac{m_i-1}{M}\le\langle\rho_i(n+k)\rangle<\frac{m_i}{M}} |\sum_{i=1}^{\mu} a_i e^{j2\pi\left(\frac{s'_i}{s}+\left\langle \mathbf{q}'_\mathbf{i},\rho'\right\rangle(n+k)\right)}|$$

$$\le \frac{8\pi\Gamma}{M}\sum_{i=1}^{\mu}\sum_{j=1}^{r}|q_{i,j}|.$$

By selecting M such that $\frac{8\pi\Gamma}{M}\sum_{i=1}^{\mu}\sum_{j=1}^{r}|q_{i,j}|\le\epsilon$, (7.51) is upper bounded by

$$(7.51)\le\sum_{n=1}^{N}\sum_{1\le m_1\le M,\cdots,1\le m_r\le M}\sum_{0\le s'<s}\mathbb{P}\{|\sum_{i=1}^{\mu} a_i e^{j2\pi\left(\frac{s'}{s}+\left\langle \mathbf{q}'_\mathbf{i},\mathbf{X}_{\mathbf{m_1},\cdots,\mathbf{m_r}}\right\rangle\right)}|<2\epsilon\}\cdot\mathbf{1}\{n\in S_{m_1,\cdots,m_r}\}. \tag{7.53}$$

Since $(\langle \rho_1 n \rangle, \cdots, \langle \rho_k n \rangle)$ satisfies Weyl's criterion, by Theorem 7.3

$$\lim_{N\to\infty} \sup_{k\in\mathbb{Z}} \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}\{n \in S_{m_1,\cdots,m_r}\} = \frac{1}{M^r}. \tag{7.54}$$

Therefore, if we let $\mathbf{X}$ be a $1 \times r$ random vector whose distribution is uniform on $[0,1)^r$, by (7.53) and (7.54)

$$\begin{aligned}
&\lim_{N\to\infty} \sup_{|a_m|\geq\gamma, |a_i|\leq\Gamma, k\in\mathbb{Z}} \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}\{|\sum_{i=1}^{\mu} a_i e^{j2\pi\omega_i(n+k)}| < \epsilon\} \\
&\leq \sup_{|a_m|\geq\gamma, |a_i|\leq\Gamma, k\in\mathbb{Z}} \sum_{1\leq m_1\leq M,\cdots,1\leq m_r\leq M} \sum_{0\leq s'<s} \mathbb{P}\{|\sum_{i=1}^{\mu} a_i e^{j2\pi\left(\frac{s'}{s} + \langle \mathbf{q}_i', \mathbf{X}_{\mathbf{m_1},\cdots,\mathbf{m_r}} \rangle\right)}| < 2\epsilon\} \cdot \frac{1}{M^r} \\
&\leq \sup_{|a_m|\geq\gamma, |a_i|\leq\Gamma, k\in\mathbb{Z}} \sum_{0\leq s'<s} \mathbb{P}\{|\sum_{i=1}^{\mu} a_i e^{j2\pi\left(\frac{s'}{s} + \langle \mathbf{q}_i', \mathbf{X} \rangle\right)}| < 2\epsilon\} (\because \text{ definitions of } \mathbf{X}_{\mathbf{m_1},\cdots,\mathbf{m_r}}, \mathbf{X}) \\
&\leq \sup_{|a_m|\geq\gamma, |a_i|\leq\Gamma} s \cdot \mathbb{P}\{|\sum_{i=1}^{\mu} a_i e^{j2\pi(\langle \mathbf{q}_i', \mathbf{X} \rangle)}| < 2\epsilon\}. (\because e^{j2\pi\frac{s'}{s}} \text{ only rotates the phase.})
\end{aligned} \tag{7.55}$$

Since $\mathbf{q}_i'$ are distinct, by Lemma 7.18, (7.55) goes to 0 as $\epsilon \downarrow 0$. □

So far, we put the restriction that $|a_i| \leq \Gamma$. However, the functions are growing as $|a_i|$ increases. Therefore, Lemma 7.19 holds even after we remove such restrictions. The proof is similar to that of Lemma 7.11.

Lemma 7.20. *Let $\omega_1, \omega_2, \cdots, \omega_\mu$ be real numbers such that $\omega_i - \omega_j \notin \mathbb{Q}$ for all $i \neq j$. Then, for strictly positive numbers γ, and any $m \in \{1, \cdots, \mu\}$*

$$\lim_{\epsilon\downarrow 0} \lim_{N\to\infty} \sup_{|a_m|\geq\gamma, a_i\in\mathbb{C}, k\in\mathbb{Z}} \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}\{|\sum_{i=1}^{\mu} a_i e^{j2\pi\omega_i(n+k)}| < \epsilon\} \to 0.$$

Proof. The proof is by induction on μ, the number of terms in the inner sum.

(i) When $\mu = 1$.

Denote a_1' as $\gamma\frac{a_1}{|a_1|}$. Then,

$$\lim_{N\to\infty} \sup_{|a_1|\geq\gamma, k\in\mathbb{Z}} \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}\{|a_1 e^{j2\pi\omega_1(n+k)}| < \epsilon\} \tag{7.56}$$

$$= \lim_{N\to\infty} \sup_{|a_1|\geq\gamma, k\in\mathbb{Z}} \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}\{|\frac{\gamma}{|a_1|} a_1 e^{j2\pi\omega_1(n+k)}| < \frac{\gamma}{|a_1|}\epsilon\}$$

$$\leq \lim_{N\to\infty} \sup_{|a_1'|=1, k\in\mathbb{Z}} \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}\{|a_1' e^{j2\pi\omega_1(n+k)}| < \epsilon\} (\because \frac{\gamma}{|a_1|} \leq 1) \tag{7.57}$$

By Lemma 7.19, (7.57) converges to 0 as $\epsilon \downarrow 0$. Thus, (7.56) converges to 0 as $\epsilon \downarrow 0$.

(ii) As an induction hypothesis, we assume the lemma is true until $\mu - 1$.

To prove the lemma for μ, it is enough to show that for all $\delta > 0$ there exists $\epsilon(\delta) > 0$ such that

$$\lim_{N\to\infty} \sup_{|a_m|\geq\gamma, k\in\mathbb{Z}} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}\{|\sum_{i=1}^{\mu} a_i e^{j2\pi\omega_i(n+k)}| < \epsilon(\delta)\} < \delta.$$

By the induction hypothesis, for all $m' \neq m$ we can find $\epsilon_{m'}(\delta) > 0$ such that

$$\lim_{N\to\infty} \sup_{|a_{m'}|\geq\gamma, k\in\mathbb{Z}} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}\{|\sum_{1\leq i\leq\mu, i\neq m} a_i e^{j2\pi\omega_i(n+k)}| < \epsilon_{m'}(\delta)\} < \delta. \tag{7.58}$$

Let $\kappa(\delta) := \min\left\{\min_{m'\neq m}\left\{\frac{\epsilon_{m'}(\delta)}{2\gamma}\right\}, 1\right\}$. By Lemma 7.19, there exists $\epsilon'(\delta) > 0$ such that

$$\lim_{N\to\infty} \sup_{|a_m|\geq\gamma, |a_i|\leq\frac{\gamma}{\kappa(\delta)}, k\in\mathbb{Z}} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}\{|\sum_{i=1}^{\mu} a_i e^{j2\pi\omega_i(n+k)}| < \epsilon'(\delta)\} < \delta. \tag{7.59}$$

Set $\epsilon(\delta) := \min\left\{\epsilon'(\delta), \min_{m'\neq m}\left\{\frac{\epsilon_{m'}(\delta)}{2}\right\}\right\}$. Then, we have

$$\begin{aligned}
&\lim_{N\to\infty} \sup_{|a_m|\geq\gamma, k\in\mathbb{Z}} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}\{|\sum_{i=1}^{\mu} a_i e^{j2\pi\omega_i(n+k)}| < \epsilon(\delta)\} \\
&\leq \lim_{N\to\infty} \max\{ \sup_{|a_m|\geq\gamma, \frac{|a_i|}{|a_m|}\leq\frac{1}{\kappa(\delta)}, k\in\mathbb{Z}} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}\{|\sum_{i=1}^{\mu} a_i e^{j2\pi\omega_i(n+k)}| < \epsilon(\delta)\}, \\
&\max_{m'\neq m} \sup_{|a_m|\geq\gamma, \frac{|a_{m'}|}{|a_m|}\geq\frac{1}{\kappa(\delta)}, k\in\mathbb{Z}} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}\{|\sum_{i=1}^{\mu} a_i e^{j2\pi\omega_i(n+k)}| < \epsilon(\delta)\}\} \\
&= \max\{ \lim_{N\to\infty} \sup_{|a_m|\geq\gamma, \frac{|a_i|}{|a_m|}\leq\frac{1}{\kappa(\delta)}, k\in\mathbb{Z}} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}\{|\sum_{i=1}^{\mu} a_i e^{j2\pi\omega_i(n+k)}| < \epsilon(\delta)\}, \\
&\max_{m'\neq m} \lim_{N\to\infty} \sup_{|a_m|\geq\gamma, \frac{|a_{m'}|}{|a_m|}\geq\frac{1}{\kappa(\delta)}, k\in\mathbb{Z}} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}\{|\sum_{i=1}^{\mu} a_i e^{j2\pi\omega_i(n+k)}| < \epsilon(\delta)\}\}.
\end{aligned} \tag{7.60}$$

Let $a'_i := \frac{\gamma}{|a_m|} a_i$. Then, the first term in (7.60) is upper bounded by

$$\begin{aligned}
&\lim_{N\to\infty} \sup_{|a_m|\geq\gamma, \frac{|a_i|}{|a_m|}\leq\frac{1}{\kappa(\delta)}, k\in\mathbb{Z}} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}\{|\sum_{i=1}^{\mu} a_i e^{j2\pi\omega_i(n+k)}| < \epsilon(\delta)\} \\
&= \lim_{N\to\infty} \sup_{|a_m|\geq\gamma, \frac{|a_i|}{|a_m|}\leq\frac{1}{\kappa(\delta)}, k\in\mathbb{Z}} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}\{|\sum_{i=1}^{\mu} \frac{\gamma}{|a_m|} a_i e^{j2\pi\omega_i(n+k)}| < \frac{\gamma}{|a_m|}\epsilon(\delta)\} \\
&= \lim_{N\to\infty} \sup_{|a'_m|=\gamma, |a'_i|\leq\frac{\gamma}{\kappa(\delta)}, k\in\mathbb{Z}} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}\{|\sum_{i=1}^{\mu} a'_i e^{j2\pi\omega_i(n+k)}| < \frac{\gamma}{|a_m|}\epsilon(\delta)\} \\
&\leq \lim_{N\to\infty} \sup_{|a'_m|=\gamma, |a'_i|\leq\frac{\gamma}{\kappa(\delta)}, k\in\mathbb{Z}} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}\{|\sum_{i=1}^{\mu} a'_i e^{j2\pi\omega_i(n+k)}| < \epsilon(\delta)\}(\because \frac{\gamma}{|a_m|} \leq 1) \\
&\leq \lim_{N\to\infty} \sup_{|a'_m|=\gamma, |a'_i|\leq\frac{\gamma}{\kappa(\delta)}, k\in\mathbb{Z}} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}\{|\sum_{i=1}^{\mu} a'_i e^{j2\pi\omega_i(n+k)}| < \epsilon'(\delta)\}(\because \epsilon' \geq \epsilon) \\
&< \delta.(\because (7.59)) \qquad (7.61)
\end{aligned}$$

Let $a''_i := \frac{\gamma}{|a_{m'}|} a_i$. Then, the second term in (7.60) is upper bounded by

$$\begin{aligned}
&\lim_{N\to\infty} \sup_{|a_m|\geq\gamma, \frac{|a_{m'}|}{|a_m|}\geq\frac{1}{\kappa(\delta)}, k\in\mathbb{Z}} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}\{|\sum_{i=1}^{\mu} a_i e^{j2\pi\omega_i(n+k)}| < \epsilon(\delta)\} \\
&= \lim_{N\to\infty} \sup_{|a_m|\geq\gamma, \frac{|a_{m'}|}{|a_m|}\geq\frac{1}{\kappa(\delta)}, k\in\mathbb{Z}} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}\{|\sum_{i=1}^{\mu} \frac{\gamma}{|a_{m'}|} a_i e^{j2\pi\omega_i(n+k)}| < \frac{\gamma}{|a_{m'}|}\epsilon(\delta)\} \\
&\leq \lim_{N\to\infty} \sup_{|a_m|\geq\gamma, \frac{|a_{m'}|}{|a_m|}\geq\frac{1}{\kappa(\delta)}, k\in\mathbb{Z}} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}\{|\sum_{i=1}^{\mu} \frac{\gamma}{|a_{m'}|} a_i e^{j2\pi\omega_i(n+k)} - \frac{\gamma}{|a_{m'}|} a_m e^{j2\pi\omega_m(n+k)}| < \frac{\gamma}{|a_{m'}|}\epsilon(\delta) + \frac{\gamma}{|a_{m'}|}|a_m|\} \\
&\leq \lim_{N\to\infty} \sup_{|a_m|\geq\gamma, \frac{|a_{m'}|}{|a_m|}\geq\frac{1}{\kappa(\delta)}, k\in\mathbb{Z}} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}\{|\sum_{i=1}^{\mu} \frac{\gamma}{|a_{m'}|} a_i e^{j2\pi\omega_i(n+k)} - \frac{\gamma}{|a_{m'}|} a_m e^{j2\pi\omega_m(n+k)}| < \epsilon_{m'}(\delta)\} \\
& \qquad (7.62) \\
&\leq \lim_{N\to\infty} \sup_{|a''_{m'}|=\gamma, k\in\mathbb{Z}} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}\{|\sum_{1\leq i\leq\mu, i\neq m} a''_i e^{j2\pi\omega_i(n+k)}| < \epsilon_{m'}(\delta)\}(\because \text{definition of } a''_i) \\
&< \delta.(\because (7.58)) \qquad (7.63)
\end{aligned}$$

Here, (7.62) is justified as follows:

$$\begin{aligned}
&\frac{\gamma}{|a'_m|}\epsilon(\delta) + \frac{\gamma}{|a'_m|}|a_m| \\
&\leq \frac{\gamma}{|a_m|}\epsilon(\delta) + \gamma\kappa(\delta)(\because \frac{|a_{m'}|}{|a_m|} \geq \frac{1}{\kappa(\delta)}, \text{ and by definition } \kappa(\delta) \leq 1) \\
&\leq \epsilon(\delta) + \gamma\kappa(\delta)(\because |a_m| \geq \gamma) \\
&\leq \frac{\epsilon_{m'}(\delta)}{2} + \frac{\epsilon_{m'}(\delta)}{2}.(\because \text{definitions of } \epsilon(\delta), \kappa(\delta))
\end{aligned}$$

Therefore, by plugging (7.61) and (7.63) into (7.60), we get

$$\lim_{N\to\infty} \sup_{|a_m|\geq\gamma, k\in\mathbb{Z}} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}\{|\sum_{i=1}^{\mu} a_i e^{j2\pi\omega_i(n+k)}| < \epsilon(\delta)\} < \delta,$$

which finishes the proof. □

Now, we will generalize Lemma 7.20 by introducing polynomial terms. First, we prove that a set of polynomials is uniformly bounded away from 0 when there is a nonzero coefficient.

Lemma 7.21. *For all $n \in \mathbb{N}$, $n' \in \mathbb{Z}^+$, $m \in \{1, \cdots, n\}$, $\gamma > 0$ and $k > 0$,*

$$\lim_{T\to\infty} \sup_{|a_m|\geq\gamma, a_i\in\mathbb{C}} \frac{|\{x \in (0,T] : |\sum_{i=-n'}^{n} a_i x^i| < k\}|_{\mathbb{L}}}{T} = 0$$

where $|\cdot|_{\mathbb{L}}$ is the Lebesgue measure of the set.

Proof. Let X be a uniform random variable on $(0,1]$. Then, we have

$$\begin{aligned}
&\sup_{|a_m|\geq\gamma} \frac{|\{x \in (0,T] : |\sum_{i=-n'}^{n} a_i x^i| < k\}|_{\mathbb{L}}}{T} \\
&= \sup_{|a_m|\geq\gamma} \frac{|\{x \in (0,T] : |\sum_{i=-n'}^{n} a_i \frac{x^i}{T^m}| < \frac{k}{T^m}\}|_{\mathbb{L}}}{T} \\
&= \sup_{|a_m|\geq\gamma} \frac{|\{x \in (0,T] : |\sum_{i=-n'}^{n} a_i \left(\frac{x}{T}\right)^i| < \frac{k}{T^m}\}|_{\mathbb{L}}}{T} \\
&= \sup_{|a_m|\geq\gamma} |\{x \in (0,1] : |\sum_{i=-n'}^{n} a_i x^i| < \frac{k}{T^m}\}|_{\mathbb{L}} \\
&= \sup_{|a_m|\geq\gamma} \mathbb{P}\{|\sum_{i=-n'}^{n} a_i X^i| < \frac{k}{T^m}\} \\
&= \sup_{|a_{m+n'}|\geq\gamma} \mathbb{P}\{|\sum_{i=0}^{n+n'} a_i X^i| < \frac{kX^{n'}}{T^m}\} \\
&\leq \sup_{|a_{m+n'}|\geq\gamma} \mathbb{P}\{|\sum_{i=0}^{n+n'} a_i X^i| < \frac{k}{T^m}\}. (\because 0 < X \leq 1 \text{ w.p. } 1)
\end{aligned}$$

Therefore, by Lemma 7.11

$$\begin{aligned}
&\lim_{T\to\infty} \sup_{|a_m|\geq\gamma, a_i\in\mathbb{C}} \frac{|\{x \in [0,T] : |\sum_{i=-n'}^{n} a_i x^i| < k\}|_{\mathbb{L}}}{T} \\
&= \lim_{T\to\infty} \sup_{|a_{m+n'}|\geq\gamma, a_i\in\mathbb{C}} \mathbb{P}\{|\sum_{i=0}^{n+n'} a_i X^i| < \frac{k}{T^m}\} = 0,
\end{aligned}$$

which finishes the proof. □

The following lemma shows that the above lemma still holds even if we change Lebesgue measure to counting measure.

Lemma 7.22. *For all $n \in \mathbb{N}$, $n' \in \mathbb{Z}^+$, $m \in \{1, \cdots, n\}$, $\gamma > 0$ and $k > 0$,*

$$\lim_{N\to\infty} \sup_{|a_m|\geq\gamma, a_i\in\mathbb{C}} \frac{|\{x \in \{1, \cdots, N\} : |\sum_{i=-n'}^{n} a_i x^i| < k\}|_{\mathbb{C}}}{N} = 0$$

where $|\cdot|_{\mathbb{C}}$ implies the counting measure of the set, the cardinality of the set.

Proof. First, we will prove the following claim which relates Lebesgue measure with counting measure.

Claim 7.4. *Let $f : \mathbb{R}^+ \to \mathbb{R}$ be a $\mathcal{C}^\infty$ function with l local maxima and minima. Then,*

$$|\{x \in [1, N] : f(x) > 0\}|_{\mathbb{L}} \leq |\{x \in \{1, \cdots, N\} : f(x) > 0\}|_{\mathbb{C}} + 3l + 2.$$

Proof. Since $f(x)$ is a continuous function with l local maxima and minima, we can prove that there exist $l' \leq l + 1$, s_i and t_i $(1 \leq i \leq l')$ such that

$$\{x \in \{1, \cdots, N\} : f(x) > 0\} = \{s_1, s_1 + 1, \cdots, s_1 + t_1\} \cup \cdots \cup \{s_{l'}, s_{l'} + 1, \cdots, s_{l'} + t_{l'}\}.$$

One way to justify this is by contradiction, i.e. if we assume $l' > l + 1$, there should exist more than l local maxima and minima by the mean value theorem. Moreover, since the number of local maxima and minima is bounded by l, we have

$$\begin{aligned}
|\{x \in [1, N] : f(x) > 0\}|_{\mathbb{L}} &\leq |[s_1 - 1, s_1 + t_1 + 1]|_{\mathbb{L}} + \cdots + |[s_{l'} - 1, s_{l'} + t_{l'} + 1]|_{\mathbb{L}} + l \\
&\leq (t_1 + 2) + \cdots + (t_{l'} + 2) + l \\
&\leq |\{x \in \{1, \cdots, N\} : f(x) > 0\}|_{\mathbb{C}} + 2l' + l \\
&\leq |\{x \in \{1, \cdots, N\} : f(x) > 0\}|_{\mathbb{C}} + 3l + 2.
\end{aligned}$$

Thus, the claim is true. □

To prove the lemma, let $a_i = a_{R,i} + j a_{I,i}$ where $a_{R,i}, a_{I,i} \in \mathbb{R}$. Then,

$$\begin{aligned}
&|\sum_{i=-n'}^{n} a_i x^i| < k \\
&(\Leftrightarrow)|\sum_{i=0}^{n+n'} a_{i-n'} x^i| < k x^{n'} \\
&(\Leftrightarrow)(\sum_{i=0}^{n+n'} a_{R,i-n'} x^i)^2 + (\sum_{i=0}^{n+n'} a_{I,i-n'} x^i)^2 < k^2 x^{n'}.
\end{aligned}$$

Since $k^2x^{2n'} - (\sum_{i=0}^{n+n'} a_{R,i-n'}x^i)^2 - (\sum_{i=0}^{n+n'} a_{I,i-n'}x^i)^2$ is a continuous function with at most $2(n+n')$ local maxima and minima, by the claim we have

$$\begin{aligned}
&\lim_{N\to\infty} \sup_{|a_m|\geq\gamma, a_i\in\mathbb{C}} \frac{|\{x \in \{1,\cdots,N\} : |\sum_{i=0}^n a_i x^i| < k\}|_{\mathbb{C}}}{N} \\
&\leq \lim_{N\to\infty} \sup_{|a_m|\geq\gamma, a_i\in\mathbb{C}} \frac{|\{x \in [1,N] : |\sum_{i=0}^n a_i x^i| < k\}|_{\mathbb{L}} + 6(n+n') + 2}{N} \\
&= \lim_{N\to\infty} \sup_{|a_m|\geq\gamma, a_i\in\mathbb{C}} \frac{|\{x \in (0,N] : |\sum_{i=0}^n a_i x^i| < k\}|_{\mathbb{L}}}{N} = 0 \; (\because \textit{Lemma } 7.21)
\end{aligned}$$

Therefore, the lemma is proved. □

Now, we merge Lemma 7.22 with Lemma 7.20 to prove that Lemma 7.23 still holds even after we introduce polynomial terms to the functions.

Lemma 7.23. *Let $\omega_1, \omega_2, \cdots, \omega_\mu$ be real numbers such that $\omega_i - \omega_j \notin \mathbb{Q}$ for all $i \neq j$. Then, for strictly positive numbers γ,*

$$\lim_{\epsilon\downarrow 0} \lim_{N\to\infty} \sup_{|a_{1\nu_1}|\geq\gamma, a_{ij}\in\mathbb{C}, k\in\mathbb{Z}} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}\left\{\left|\sum_{i=1}^{\mu}\left(\sum_{j=0}^{\nu_i} a_{ij}n^j\right) e^{j2\pi\omega_i(n+k)}\right| < \epsilon\right\} \to 0.$$

Proof. To prove the lemma, it is enough to show that for all $\delta > 0$, there exist $\epsilon > 0$ and $N_T \in \mathbb{N}$ such that for all $N \geq N_T$,

$$\sup_{|a_{1\nu_1}|\geq\gamma, a_{ij}\in\mathbb{C}, k\in\mathbb{Z}} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}\left\{\left|\sum_{i=1}^{\mu}\left(\sum_{j=0}^{\nu_i} a_{ij}n^j\right) e^{j2\pi\omega_i(n+k)}\right| < \epsilon\right\} < \delta. \tag{7.64}$$

Since μ is finite, by Lemma 7.20, there exist $\epsilon' > 0$ and $M_T \in \mathbb{N}$ such that for all $M \geq M_T$,

$$\max_{d\in\{1,\cdots,\mu\}} \left(\sup_{k\in\mathbb{Z}, a_i\in\mathbb{C}, |a_d|\geq 1} \frac{1}{M}\sum_{c=1}^{M} \mathbf{1}\left\{\left|\sum_{i=1}^{\mu} a_i e^{j2\pi\omega_i(c+k)}\right| < \epsilon'\right\}\right) < \frac{\delta}{4}. \tag{7.65}$$

By Lemma 7.22, there exists $B'_T \in \mathbb{N}$ such that for all $B' \geq B'_T$,

$$\sup_{|a'_{1\nu_1}|\geq\gamma} \frac{\left|\left\{b \in \{1,\cdots,B'\} : \left|\sum_{j=0}^{\nu_1} a'_{1j}b^j\right| \leq 2\right\}\right|_{\mathbb{C}}}{B'} < \frac{\delta}{4}. \tag{7.66}$$

Define $\kappa' := \frac{2\sum_{i=1}^{\mu}\sum_{j=1}^{\nu_i}\sum_{k=1}^{j}\binom{j}{k}}{\epsilon'}$. By Lemma 7.22, there exists $B''_T \in \mathbb{N}$ such that for all $B'' \geq B''_T$,

$$\sum_{1\leq i\leq\mu, 1\leq j\leq\nu_i, 1\leq k\leq j} \sup_{|a_k|=1} \frac{|\{b \in \{1,\cdots,B''\} : \kappa' \geq |\sum_{j'=-j+k}^{\nu_i-j+k} a_{j'}b^{j'}|\}|_{\mathbb{C}}}{B''} < \frac{\delta}{4}. \tag{7.67}$$

Define $B := \max(B'_T, B''_T)$. We will show that the choice of $\epsilon = \epsilon'$ and $N_T = \max(M_T, \lceil \frac{4}{\delta} \rceil) \cdot B$ satisfies (7.64). Then, for all $N \geq N_T$, N can be written as $N = B \cdot M + b$ for some $M \geq \max(M_T, \lceil \frac{4}{\delta} \rceil)$ and $0 \leq b \leq B-1$.

$$
\begin{aligned}
&\sup_{|a_{1\nu_1}|\geq\gamma, a_{ij}\in\mathbb{C}, k\in\mathbb{Z}} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}\left\{\left|\sum_{i=1}^{\mu}\left(\sum_{j=0}^{\nu_i} a_{ij} n^j\right) e^{j2\pi\omega_i(n+k)}\right| < \epsilon\right\} \\
&= \sup_{|a_{1\nu_1}|\geq\gamma, a_{ij}\in\mathbb{C}} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}\left\{\left|\sum_{i=1}^{\mu}\left(\sum_{j=0}^{\nu_i} a_{ij} n^j\right) e^{j2\pi\omega_i n}\right| < \epsilon\right\} (\because e^{j2\pi\omega_i k} \text{ can be absorbed into the } a_{ij}.) \\
&= \sup_{|a_{1\nu_1}|\geq\gamma, a_{ij}\in\mathbb{C}} \frac{1}{B\cdot M+b}\sum_{n=1}^{B\cdot M+b} \mathbf{1}\left\{\left|\sum_{i=1}^{\mu}\left(\sum_{j=0}^{\nu_i} a_{ij} n^j\right) e^{j2\pi\omega_i n}\right| < \epsilon\right\} \\
&\leq \sup_{|a_{1\nu_1}|\geq\gamma, a_{ij}\in\mathbb{C}} \frac{1}{B\cdot M}\sum_{n=1}^{B\cdot M} \mathbf{1}\left\{\left|\sum_{i=1}^{\mu}\left(\sum_{j=0}^{\nu_i} a_{ij} n^j\right) e^{j2\pi\omega_i n}\right| < \epsilon\right\} + \frac{b}{B\cdot M} \\
&\leq \sup_{|a_{1\nu_1}|\geq\gamma, a_{ij}\in\mathbb{C}} \frac{1}{B\cdot M}\sum_{n=1}^{B\cdot M} \mathbf{1}\left\{\left|\sum_{i=1}^{\mu}\left(\sum_{j=0}^{\nu_i} a_{ij} n^j\right) e^{j2\pi\omega_i n}\right| < \epsilon\right\} + \frac{\delta}{4} \\
&= \sup_{|a_{1\nu_1}|\geq\gamma, a_{ij}\in\mathbb{C}} \frac{1}{B\cdot M}\sum_{b=0}^{B-1}\sum_{c=1}^{M} \mathbf{1}\left\{\left|\sum_{i=1}^{\mu}\left(\sum_{j=0}^{\nu_i} a_{ij} (bM+c)^j\right) e^{j2\pi\omega_i (bM+c)}\right| < \epsilon\right\} + \frac{\delta}{4} (\because n \text{ is rewritten as } bM+c.) \\
&= \sup_{|a_{1\nu_1}|\geq\gamma, a_{ij}\in\mathbb{C}} \frac{1}{B\cdot M}\sum_{b=0}^{B-1}\sum_{c=1}^{M} \mathbf{1}\left\{\left|\sum_{i=1}^{\mu}\left(\sum_{j=0}^{\nu_i} a_{ij} \left((bM)^j + \sum_{k=1}^{j}\binom{j}{k}(bM)^{j-k}c^k\right)\right) e^{j2\pi\omega_i (bM+c)}\right| < \epsilon\right\} + \frac{\delta}{4} \\
&\leq \sup_{|a_{1\nu_1}|\geq\gamma, a_{ij}\in\mathbb{C}} \frac{1}{B\cdot M}\sum_{b=0}^{B-1}\sum_{c=1}^{M} \mathbf{1}\left\{\left|\sum_{i=1}^{\mu}\left(\sum_{j=0}^{\nu_i} a_{ij} (bM)^j\right) e^{j2\pi\omega_i (bM+c)}\right| < \epsilon + \sum_{i=1}^{\mu}\sum_{j=1}^{\nu_i}|a_{ij}|\sum_{k=1}^{j}\binom{j}{k}(bM)^{j-k}c^k\right\} + \frac{\delta}{4} \\
&\leq \sup_{|a_{1\nu_1}|\geq\gamma, a_{ij}\in\mathbb{C}} \frac{1}{B\cdot M}\sum_{b=0}^{B-1}\sum_{c=1}^{M} \mathbf{1}\left\{\left|\sum_{i=1}^{\mu}\left(\sum_{j=0}^{\nu_i} a_{ij} (bM)^j\right) e^{j2\pi\omega_i (bM+c)}\right| < \epsilon + \sum_{i=1}^{\mu}\sum_{j=1}^{\nu_i}\sum_{k=1}^{j}|a_{ij}|\binom{j}{k}(bM)^{j-k}M^k\right\} + \frac{\delta}{4}
\end{aligned}
\tag{7.68}
$$

$$
\begin{aligned}
&\leq \sup_{|a_{1\nu_1}|\geq\gamma, a_{ij}\in\mathbb{C}} \frac{1}{B\cdot M}\sum_{b=0}^{B-1}\sum_{c=1}^{M} \mathbf{1}\Bigg\{\left|\sum_{i=1}^{\mu}\left(\frac{\sum_{j=0}^{\nu_i} a_{ij}(bM)^j}{M_b}\right) e^{j2\pi\omega_i (bM+c)}\right| < \\
&\frac{\epsilon}{M_b} + \frac{\sum_{i=1}^{\mu}\sum_{j=1}^{\nu_i}\sum_{k=1}^{j}|a_{ij}|\binom{j}{k}(bM)^{j-k}M^k}{M_b}\Bigg\} + \frac{\delta}{4} \\
&\leq \sup_{|a_{1\nu_1}|\geq\gamma, a_{ij}\in\mathbb{C}} \frac{1}{B}\sum_{b=0}^{B-1}\Bigg\{\frac{1}{M}\sum_{c=1}^{M}\mathbf{1}\left\{\left|\sum_{i=1}^{\mu}\left(\frac{\sum_{j=0}^{\nu_i} a_{ij}(bM)^j}{M_b}\right) e^{j2\pi\omega_i (bM+c)}\right| < \epsilon\right\} \\
&+\mathbf{1}\left\{\frac{\epsilon}{M_b} + \frac{\sum_{i=1}^{\mu}\sum_{j=1}^{\nu_i}\sum_{k=1}^{j}|a_{ij}|\binom{j}{k}(bM)^{j-k}M^k}{M_b} \geq \epsilon\right\}\Bigg\} + \frac{\delta}{4} \\
&\leq \sup_{|a_{1\nu_1}|\geq\gamma, a_{ij}\in\mathbb{C}} \frac{1}{B}\sum_{b=0}^{B-1}\Bigg\{\frac{1}{M}\sum_{c=1}^{M}\mathbf{1}\left\{\left|\sum_{i=1}^{\mu}\left(\frac{\sum_{j=0}^{\nu_i} a_{ij}(bM)^j}{M_b}\right) e^{j2\pi\omega_i (bM+c)}\right| < \epsilon\right\} \\
&+\mathbf{1}\left\{\frac{\epsilon}{M_b} \geq \frac{\epsilon}{2}\right\} + \mathbf{1}\left\{\frac{\sum_{i=1}^{\mu}\sum_{j=1}^{\nu_i}\sum_{k=1}^{j}|a_{ij}|\binom{j}{k}(bM)^{j-k}M^k}{M_b} \geq \frac{\epsilon}{2}\right\}\Bigg\} + \frac{\delta}{4}
\end{aligned}
\tag{7.69}
$$

where $M_b := \max_i \left\{\left|\sum_{j=0}^{\nu_i} a_{ij}\,(bM)^j\right|\right\}$ and when $M_b = 0$ the value of the indicator function is set to be 0 since in this case, the indicator function of (7.68) is already 0.

First, let's prove that the first term of (7.69) is small enough. For all $a_{ij} \in \mathbb{C}$ such that $|a_{1\nu_1}| \geq \gamma$

and $b \in \{0, \cdots, B\}$, we have

$$\begin{aligned}
&\frac{1}{M}\sum_{c=1}^{M}\mathbf{1}\left\{\left|\sum_{i=1}^{\mu}\left(\frac{\sum_{j=0}^{\nu_i} a_{ij}(bM)^j}{M_b}\right)e^{j2\pi\omega_i(bM+c)}\right| < \epsilon\right\} \\
&\le \max_{d\in\{1,\cdots,\mu\}}\left(\sup_{k\in\mathbb{Z},a_i\in\mathbb{C},|a_d|\ge 1}\frac{1}{M}\sum_{c=1}^{M}\mathbf{1}\left\{\left|\sum_{i=1}^{\mu}a_i e^{j\omega_i(c+k)}\right| < \epsilon\right\}\right) \\
&(\because \text{ By the definition of } M_b, \left|\frac{\sum_{j=0}^{\nu_i} a_{ij}(bM)^j}{M_b}\right| = 1 \text{ for some } i) \\
&< \frac{\delta}{4}.(\because (7.65))
\end{aligned} \tag{7.70}$$

Let's prove that the second term of (7.69) is small enough.

$$\begin{aligned}
&\sup_{|a_{1\nu_1}|\ge\gamma}\frac{|\{b\in\{1,\cdots,B\}: M_b < 2\}|_{\mathbb{C}}}{B} \\
&\le \sup_{|a_{1\nu_1}|\ge\gamma}\frac{\left|\left\{b\in\{1,\cdots,B\}:\left|\sum_{j=0}^{\nu_1}a_{1j}(bM)^j\right| < 2\right\}\right|_{\mathbb{C}}}{B}(\because \text{ definition of } M_b) \\
&\le \sup_{|a'_{1\nu_1}|\ge\gamma}\frac{\left|\left\{b\in\{1,\cdots,B\}:\left|\sum_{j=0}^{\nu_1}a'_{1j}b^j\right| < 2\right\}\right|_{\mathbb{C}}}{B}(\because \text{ putting } a'_{1j} := a_{1j}M^j \text{ and } M \text{ goes to infinity.}) \\
&< \frac{\delta}{4}.(\because (7.66))
\end{aligned} \tag{7.71}$$

Now, we will prove that the third term of (7.69) is small enough.

$$\begin{aligned}
&\sup_{|a_{1\nu_1}|\ge\gamma,a_{ij}\in\mathbb{C}}\frac{1}{B}\sum_{b=0}^{B-1}\mathbf{1}\left\{\frac{\sum_{i=1}^{\mu}\sum_{j=1}^{\nu_i}\sum_{k=1}^{j}|a_{ij}|\binom{j}{k}(bM)^{j-k}M^k}{M_b}\ge\frac{\epsilon}{2}\right\} \\
&\le \sup_{|a_{1\nu_1}|\ge\gamma,a_{ij}\in\mathbb{C}}\frac{1}{B}\sum_{b=0}^{B-1}\mathbf{1}\left\{\left(\sum_{i'=1}^{\mu}\sum_{j'=1}^{\nu_{i'}}\sum_{k'=1}^{j'}\binom{j'}{k'}\right)\cdot\max_{1\le i\le\mu,1\le j\le\nu_i,1\le k\le j}|a_{ij}|(bM)^{j-k}M^k\ge\frac{\epsilon}{2}M_b\right\} \\
&\le \sum_{1\le i\le\mu,1\le j\le\nu_i,1\le k\le j}\sup_{|a_{1\nu_1}|\ge\gamma,a_{ij}\in\mathbb{C}}\frac{1}{B}\sum_{b=0}^{B-1}\mathbf{1}\left\{\kappa'|a_{ij}|(bM)^{j-k}M^k\ge M_b\right\} \\
&\le \sum_{1\le i\le\mu,1\le j\le\nu_i,1\le k\le j}\sup_{|a_{1\nu_1}|\ge\gamma,a_{ij}\in\mathbb{C}}\frac{1}{B}\sum_{b=0}^{B-1}\mathbf{1}\left\{\kappa'|a_{ij}|(bM)^{j-k}M^k\ge\left|\sum_{j'=0}^{\nu_i}a_{ij'}(bM)^{j'}\right|\right\}(\because \text{ definition of } M_b) \\
&\le \sum_{1\le i\le\mu,1\le j\le\nu_i,1\le k\le j}\sup_{|a_{1\nu_1}|\ge\gamma,a_{ij}\in\mathbb{C}}\frac{1}{B}\sum_{b=0}^{B-1}\mathbf{1}\left\{\kappa'\ge\left|\sum_{j'=0}^{\nu_i}\frac{a_{ij'}(bM)^{j'}}{|a_{ij}|b^{j-k}M^k}\right|\right\} \\
&\le \sum_{1\le i\le\mu,1\le j\le\nu_i,1\le k\le j}\sup_{|a_k|=1}\frac{1}{B}\sum_{b=0}^{B-1}\mathbf{1}\left\{\kappa'\ge\left|\sum_{j'=-j+k}^{\nu_i-j+k}a_{j'}b^{j'}\right|\right\} \\
&< \frac{\delta}{4}.(\because (7.67))
\end{aligned} \tag{7.72}$$

Therefore, by (7.70), (7.71), (7.72), we can see (7.69) $< \delta$, which finishes the proof. □

7.7 Proof of Lemma 2.3

In this section, we will merge the properties about the observability Gramian shown in Appendix 7.5 with the uniform convergence of Appendix 7.6, and prove Lemma 2.3 of page 63.

Just as we did in Appendix 7.4, we must first prove the following lemma which tells us that the determinant of the observability Gramian is large except on a negligible set under a cofactor condition the Gramian matrix. The proof of the lemma is very similar to that of Lemma 7.12.

Lemma 7.24. *Let* $\mathbf{A}$ *and* $\mathbf{C}$ *be given as* (7.39) *and* (7.40). *Define* $a_{i,j}$ *and* $C_{i,j}$ *as the* (i,j) *element and cofactor of* $\begin{bmatrix} \mathbf{CA}^{-k_1} \\ \vdots \\ \mathbf{CA}^{-k_{m-1}} \\ \mathbf{CA}^{-n} \end{bmatrix}$ *respectively. Then, there exists a family of functions* $\{g_\epsilon : \epsilon > 0, g_\epsilon : \mathbb{R}^+ \to \mathbb{R}^+\}$ *satisfying:*

(i) For all $\epsilon > 0$, $k_1 < k_2 < \cdots < k_{m-1}$ *and* $|C_{m,m}| \geq \epsilon \prod_{1 \leq i \leq m-1} \lambda_i^{-k_i}$, *the following is true.*

$$\lim_{N \to \infty} \sup_{k \in \mathbb{Z}, k - k_{m-1} \geq g_\epsilon(k_{m-1})} \frac{1}{N} \sum_{n=k+1}^{k+N} \mathbf{1}\left\{ \left| \det \left(\begin{bmatrix} \mathbf{CA}^{-k_1} \\ \vdots \\ \mathbf{CA}^{-k_{m-1}} \\ \mathbf{CA}^{-n} \end{bmatrix} \right) \right| < \epsilon^2 \lambda_m^{-n} \prod_{1 \leq i \leq m-1} \lambda_i^{-k_i} \right\} \to 0 \text{ as } \epsilon \downarrow 0.$$

(ii) For each $\epsilon > 0$, $g_\epsilon(k) \lesssim 1 + \log(k+1)$.

Proof. By Lemma 7.16, we can find a function $g'_{2\epsilon^2}(k)$ such that for all $0 \leq k_1 < k_2 < \cdots < k_{m-1} < n$ satisfying:

(i) $n - k_{m-1} \geq g'_{2\epsilon^2}(k_{m-1})$

(ii) $g'_{2\epsilon^2}(k) \lesssim 1 + \log(k+1)$

(iii) $\left| \sum_{m - m_\mu + 1 \leq i \leq m} a_{m,i} C_{m,i} \right| \geq 2\epsilon^2 \lambda_m^{-n} \prod_{1 \leq i \leq m-1} \lambda_i^{-k_i}$

the following inequality holds:

$$\left| \det \left(\begin{bmatrix} \mathbf{CA}^{-k_1} \\ \vdots \\ \mathbf{CA}^{-k_{m-1}} \\ \mathbf{CA}^{-n} \end{bmatrix} \right) \right| \geq \epsilon^2 \lambda_m^{-n} \prod_{1 \leq i \leq m-1} \lambda_i^{-k_i}.$$

Let $g_\epsilon(k)$ be $g'_{2\epsilon^2}(k)$. Then, we have

$$\sup_{k\in\mathbb{Z},k-k_{m-1}\geq g_\epsilon(k_{m-1})} \frac{1}{N}\sum_{n=k+1}^{k+N} \mathbf{1}\left\{\left|\det\left(\begin{bmatrix}\mathbf{CA}^{-k_1}\\ \vdots \\ \mathbf{CA}^{-k_{m-1}} \\ \mathbf{CA}^{-n}\end{bmatrix}\right)\right| < \epsilon^2\lambda_m^{-n}\prod_{1\leq i\leq m-1}\lambda_i^{-k_i}\right\}$$

$$\leq \sup_{k\in\mathbb{Z},k-k_{m-1}\geq g_\epsilon(k_{m-1})} \frac{1}{N}\sum_{n=k+1}^{k+N} \mathbf{1}\left\{\left|\sum_{m-m_\mu+1\leq i\leq m} a_{m,i}C_{m,i}\right| < 2\epsilon^2\lambda_m^{-n}\prod_{1\leq i\leq m-1}\lambda_i^{-k_i}\right\} \quad (7.73)$$

$$= \sup_{k\in\mathbb{Z},k-k_{m-1}\geq g_\epsilon(k_{m-1})} \frac{1}{N}\sum_{n=k+1}^{k+N} \mathbf{1}\left\{\left|\sum_{m-m_\mu+1\leq i\leq m} \frac{a_{m,i}}{\lambda_m^{-n}}\frac{C_{m,i}}{\epsilon\prod_{1\leq i\leq m-1}\lambda_i^{-k_i}}\right| < 2\epsilon\right\}$$

$$\leq \sup_{k\in\mathbb{Z},|b_m|\geq 1} \frac{1}{N}\sum_{n=k+1}^{k+N} \mathbf{1}\left\{\left|\sum_{m-m_\mu+1\leq i\leq m} b_i\frac{a_{m,i}}{\lambda_m^{-n}}\right| < 2\epsilon\right\} \quad (7.74)$$

where (7.73) is by the definition of $g_\epsilon(k)$ and Lemma 7.16, and (7.74) is by $|C_{m,m}| \geq \epsilon\prod_{1\leq i\leq m-1}\lambda_i^{-k_i}$.

Let $\mathbf{C}_{\mu,\nu_\mu}$ denoted in (7.40) be $\begin{bmatrix}c'_1 & \cdots & c'_{m_{\mu,\nu_\mu}}\end{bmatrix}$.

Moreover,

$$\mathbf{A}_{\mu,\nu_\mu}{}^{-n}$$
$$= \begin{bmatrix} (\lambda_{\mu,\nu_\mu}e^{j2\pi\omega_{\mu,\nu_\mu}})^{-n} & \binom{-n}{1}(\lambda_{\mu,\nu_\mu}e^{j2\pi\omega_{\mu,\nu_\mu}})^{-n-1} & \cdots & \binom{-n}{m_{\mu,\nu_\mu}-1}(\lambda_{\mu,\nu_\mu}e^{j2\pi\omega_{\mu,\nu_\mu}})^{-n-m_{\mu,\nu_\mu}+1} \\ 0 & (\lambda_{\mu,\nu_\mu}e^{j2\pi\omega_{\mu,\nu_\mu}})^{-n} & \cdots & \binom{-n}{m_{\mu,\nu_\mu}-2}(\lambda_{\mu,\nu_\mu}e^{j2\pi\omega_{\mu,\nu_\mu}})^{-n-m_{\mu,\nu_\mu}+2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & (\lambda_{\mu,\nu_\mu}e^{j2\pi\omega_{\mu,\nu_\mu}})^{-n} \end{bmatrix}.$$

Thus, we can see that

$$a_{m,m} = \sum_{1\leq i\leq m_{\mu,\nu_\mu}} c'_i\binom{-n}{m_{\mu,\nu_\mu}-i}(\lambda_{\mu,\nu_\mu}e^{j2\pi\omega_{\mu,\nu_\mu}})^{-n-m_{\mu,\nu_\mu}+i}.$$

Therefore,

$$\frac{a_{m,m}}{\lambda_m^{-n}} = \sum_{1\leq i\leq m_{\mu,\nu_\mu}} c'_i\binom{-n}{m_{\mu,\nu_\mu}-i}\lambda_{\mu,\nu_\mu}^{-m_{\mu,\nu_\mu}+i}(e^{j2\pi\omega_{\mu,\nu_\mu}})^{-n-m_{\mu,\nu_\mu}+i}.$$

Moreover, when $a_{m,i}$ is considered as a function of n, the $n^{m_{\mu,\nu_\mu}-1}e^{-j2\pi\omega_{\mu,\nu_\mu}n}$ term only shows up in $\frac{a_{m,m}}{\lambda_m^{-n}}$ among $\frac{a_{m,m-m_\mu+1}}{\lambda_m^{-n}}, \cdots, \frac{a_{m,m}}{\lambda_m^{-n}}$, and the associated coefficient is $\frac{c'_1(-1)^{m_{\mu,\nu_\mu}-1}}{(m_{\mu,\nu_\mu}-1)!}\lambda_{\mu,\nu_\mu}^{-m_{\mu,\nu_\mu}+1}e^{j2\pi\omega_{\mu,\nu_\mu}(-m_{\mu,\nu_\mu}+1)}$.

Let $c' := \frac{|c_1'|}{(m_{\mu,\nu_\mu}-1)!}\lambda_{\mu,\nu_\mu}^{-m_{\mu,\nu_\mu}+1}$. Then, (7.74) can be upper bounded as follows:

$$
\begin{aligned}
(7.74) &\leq \sup_{k\in\mathbb{Z},|a_{\nu_\mu,m_{\mu,\nu_\mu}}|\geq c'} \frac{1}{N}\sum_{n=k+1}^{k+N} \mathbf{1}\left\{\left|\sum_{1\leq i\leq \nu_\mu}\left(\sum_{1\leq j\leq m_{\mu,i}} a_{ij}n^{j-1}\right)e^{j2\pi(-\omega_{\mu,i})n}\right| < 2\epsilon\right\} \\
&= \sup_{k\in\mathbb{Z},|a_{\nu_\mu,m_{\mu,\nu_\mu}}|\geq c'} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}\left\{\left|\sum_{1\leq i\leq \nu_\mu}\left(\sum_{1\leq j\leq m_{\mu,i}} a_{ij}(n+k)^{j-1}\right)e^{j2\pi(-\omega_{\mu,i})(n+k)}\right| < 2\epsilon\right\} \\
&\leq \sup_{k\in\mathbb{Z},|a_{\nu_\mu,m_{\mu,\nu_\mu}}|\geq c'} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}\left\{\left|\sum_{1\leq i\leq \nu_\mu}\left(\sum_{1\leq j\leq m_{\mu,i}} a_{ij}n^{j-1}\right)e^{j2\pi(-\omega_{\mu,i})(n+k)}\right| < 2\epsilon\right\} \quad (7.75)
\end{aligned}
$$

The last inequality comes from the fact that the coefficient of $n^{m_{\mu,\nu_\mu}-1}$ is the same for both $\sum_{1\leq j\leq m_{\mu,\nu_\mu}} a_{\nu_\mu,j}(n+k)^{j-1}$ and $\sum_{1\leq j\leq m_{\mu,\nu_\mu}} a_{\nu_\mu,j}n^{j-1}$.

By Lemma 7.11, we get

$$
\lim_{N\to\infty} \sup_{k\in\mathbb{Z},|a_{\nu_\mu,m_{\mu,\nu_\mu}}|\geq c'} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}\left\{\left|\sum_{1\leq i\leq \nu_\mu}\left(\sum_{1\leq j\leq m_{\mu,i}} a_{ij}n^{j-1}\right)e^{j2\pi(-\omega_{\mu,i})(n+k)}\right| < 2\epsilon\right\} \to 0 \text{ as } \epsilon\downarrow 0.
$$

Therefore, by (7.75) we can say that

$$
\lim_{N\to\infty} \sup_{k\in\mathbb{Z},k-k_{m-1}\geq g_\epsilon(k_{m-1})} \frac{1}{N}\sum_{n=k+1}^{k+N} \mathbf{1}\left\{\left|\det\left(\begin{bmatrix}\mathbf{CA}^{-k_1}\\ \vdots \\ \mathbf{CA}^{-k_{m-1}} \\ \mathbf{CA}^{-n}\end{bmatrix}\right)\right| < \epsilon^2\lambda_m^{-n}\prod_{1\leq i\leq m-1}\lambda_i^{-k_i}\right\} \to 0 \text{ as } \epsilon\downarrow 0
$$

which finishes the proof. □

Based on the previous lemma, the properties of p.m.f. tails shown in Section 7.1 and the properties of the observability Gramian discussed in Section 7.5, we can prove Lemma 2.3 for the case when the system has no eigenvalue cycles. Moreover, we will prove a lemma involving multiple systems. This will turn out to be helpful in proving Lemma 2.3 for general systems with eigenvalue cycles.

Consider pairs of matrices $(\mathbf{A_1},\mathbf{C_1}),(\mathbf{A_2},\mathbf{C_2}),\cdots,(\mathbf{A_r},\mathbf{C_r})$ defined as follows:

$\mathbf{A_i}$ is a $m_i\times m_i$ Jordan form matrix and $\mathbf{C_i}$ is a $1\times m_i$ row vector (7.76)

Each $\mathbf{A_i}$ has no eigenvalues cycles and $(\mathbf{A_i},\mathbf{C_i})$ is observable

$\lambda_j^{(i)}e^{j2\pi\omega_j^{(i)}}$ is (j,j) element of $\mathbf{A_i}$

$\lambda_1^{(i)}\geq\lambda_2^{(i)}\geq\cdots\geq\lambda_{m_i}^{(i)}\geq 1.$

Then, the following lemma holds.

Lemma 7.25. *Consider systems* $(\mathbf{A_1},\mathbf{C_1}),(\mathbf{A_2},\mathbf{C_2}),\cdots,(\mathbf{A_r},\mathbf{C_r})$ *given as* (7.76). *Then, we can find a polynomial* $p(k)$ *and a family of random variables* $\{S(\epsilon,k):k\in\mathbb{Z}^+,\epsilon>0\}$ *such that for all*

$\epsilon > 0$, $k \in \mathbb{Z}^+$ and $1 \leq i \leq r$ there exist $k \leq k_{i,1} < k_{i,2} < \cdots < k_{i,m_i} \leq S(\epsilon, k)$ and $\mathbf{M_i}$ satisfying the following conditions:

(i) $\beta[k_{i,j}] = 1$ *for* $1 \leq i \leq r$ *and* $1 \leq j \leq m_i$

(ii) $\mathbf{M_i} \begin{bmatrix} \mathbf{C_i}\mathbf{A_i}^{-k_{i,1}} \\ \mathbf{C_i}\mathbf{A_i}^{-k_{i,2}} \\ \vdots \\ \mathbf{C_i}\mathbf{A_i}^{-k_{i,m_i}} \end{bmatrix} = \mathbf{I}$

(iii) $|\mathbf{M_i}|_{max} \leq \frac{p(S(\epsilon,k))}{\epsilon}(\lambda_1^{(i)})^{S(\epsilon,k)}$

(iv) $\lim_{\epsilon \downarrow 0} \exp \limsup_{s \to \infty} \sup_{k \in \mathbb{Z}^+} \frac{1}{s} \log \mathbb{P}\{S(\epsilon, k) - k = s\} = p_e$.

Proof. By Lemma 7.15, instead of the conditions (ii) and (iii) it is enough to prove that

$$\left| \det \left(\begin{bmatrix} \mathbf{C_i}\mathbf{A_i}^{-k_{i,1}} \\ \mathbf{C_i}\mathbf{A_i}^{-k_{i,2}} \\ \vdots \\ \mathbf{C_i}\mathbf{A_i}^{-k_{i,m_i}} \end{bmatrix} \right) \right| \geq \epsilon \prod_{1 \leq j \leq m_i} (\lambda_j^{(i)})^{-k_{i,j}}.$$

Therefore, it is enough to prove the following claim:

Claim 7.5. *We can find a family of stopping times $\{S(\epsilon, k) : k \in \mathbb{Z}^+, \epsilon > 0\}$ such that for all $\epsilon > 0$, $k \in \mathbb{Z}^+$ and $1 \leq i \leq r$ there exist $k \leq k_{i,1} < k_{i,2} < \cdots < k_{i,m_i} \leq S(\epsilon, k)$ satisfying the following condition:*

(a) $\beta[k_{i,j}] = 1$ *for* $1 \leq i \leq r$ *and* $1 \leq j \leq m_i$

(b) $\left| \det \left(\begin{bmatrix} \mathbf{C_i}\mathbf{A_i}^{-k_{i,1}} \\ \mathbf{C_i}\mathbf{A_i}^{-k_{i,2}} \\ \vdots \\ \mathbf{C_i}\mathbf{A_i}^{-k_{i,m_i}} \end{bmatrix} \right) \right| \geq \epsilon \prod_{1 \leq j \leq m_i} (\lambda_j^{(i)})^{-k_{i,j}}$

(c) $\lim_{\epsilon \downarrow 0} \exp \limsup_{s \to \infty} \sup_{k \in \mathbb{Z}^+} \frac{1}{s} \log \mathbb{P}\{S(\epsilon, k) - k = s\} \leq p_e$.

Before we prove the above claim, we first prove the claim for a single system.

Claim 7.6. *We can find a family of stopping times $\{S_1(\epsilon, k) : k \in \mathbb{Z}^+, \epsilon > 0\}$ such that for all $\epsilon > 0$ and $k \in \mathbb{Z}^+$ there exist $k \leq k'_1 < k'_2 < \cdots < k'_{m_1} \leq S_1(\epsilon, k)$ satisfying the following condition:*

(a') $\beta[k'_j] = 1$ *for* $1 \leq j \leq m_1$

(b') $\left| \det \left(\begin{bmatrix} \mathbf{C_1}\mathbf{A_1}^{-k'_1} \\ \mathbf{C_1}\mathbf{A_1}^{-k'_2} \\ \vdots \\ \mathbf{C_1}\mathbf{A_1}^{-k'_{m_i}} \end{bmatrix} \right) \right| \geq \epsilon \prod_{1 \leq j \leq m_1} (\lambda_j^{(1)})^{-k'_j}$

(c') $\lim_{\epsilon \downarrow 0} \exp \limsup_{s \to \infty} \sup_{k \in \mathbb{Z}^+} \frac{1}{s} \log \mathbb{P}\{S_1(\epsilon, k) - k = s\} \leq p_e$.

- Proof of Claim 7.6: The proof of Claim 7.6 is an induction on m_1, the dimension of $\mathbf{A_1}$.

(i) First consider the case $m_1 = 1$.

In this case, $\mathbf{A_1}$ and $\mathbf{C_1}$ is scalar, so denote $\mathbf{A_1} := \lambda_1^{(1)} e^{j2\pi\omega_1^{(1)}}$ and $\mathbf{C_1} := c_1$. Since we only care about small enough ϵ, let $\epsilon \leq |c_1|$. Denote $S_1(\epsilon, k) := \inf\{n \geq k : \beta[n] = 1\}$ and $k_1' = S_1(\epsilon, k)$. Then, $\beta[k_1'] = 1$ and $\left|\det\left(\left[c_1(\lambda_1^{(1)} e^{j2\pi\omega_1^{(1)}})^{-k_1'}\right]\right)\right| = |c_1|(\lambda_1^{(1)})^{-k_1'} \geq \epsilon(\lambda_1^{(1)})^{-k_1'}$. Moreover, since $S_1(\epsilon, k) - k$ is a geometric random variable with probability $1 - p_e$,

$$\exp \limsup_{s\to\infty} \sup_{k \in \mathbb{Z}^+} \log \mathbb{P}\left\{S_1(\epsilon, k) - k = s\right\} = p_e.$$

Therefore, $S_1(\epsilon, k)$ satisfies all the conditions of the claim.

(ii) As an induction hypothesis, we assume the claim is true for $m_1 - 1$ and prove the claim hold for m_1.

First, we will fix $k = 0$, then we will consider general $k \in \mathbb{Z}^+$.

Denote $\mathbf{A_1'}$ be a $(m_1 - 1) \times (m_1 - 1)$ matrix obtained by removing m_1th row and column of $\mathbf{A_1}$. Likewise, $\mathbf{C_1'}$ is a $1 \times (m_1 - 1)$ vector obtained by removing m_1th element of $\mathbf{C_1}$. Then, we can observe that

$$\det\left(\begin{bmatrix} \mathbf{C_1'}\mathbf{A_1'}^{-k_1'} \\ \vdots \\ \mathbf{C_1'}\mathbf{A_1'}^{-k_{m_1-1}'} \end{bmatrix}\right) = cof_{m_1,m_1}\left(\begin{bmatrix} \mathbf{C_1}\mathbf{A_1}^{-k_1'} \\ \vdots \\ \mathbf{C_1}\mathbf{A_1}^{-k_{m_1}'} \end{bmatrix}\right)$$

where $cof_{i,j}(\mathbf{A})$ implies the cofactor matrix of $\mathbf{A}$ with respect to (i,j) element.

By the induction hypothesis, we can find a stopping time $S_1'(\epsilon, 0)$ such that there exist $0 \leq k_1' < k_2' < \cdots < k_{m_1-1}' \leq S_1'(\epsilon, 0)$ satisfying:

(a") $\beta[k_j'] = 1$ for $1 \leq j \leq m_1 - 1$

(b") $\left|\det\left(\begin{bmatrix} \mathbf{C_1'}\mathbf{A_1'}^{-k_1'} \\ \vdots \\ \mathbf{C_1'}\mathbf{A_1'}^{-k_{m_1-1}'} \end{bmatrix}\right)\right| \geq \epsilon \prod_{1 \leq j \leq m_1-1} (\lambda_j^{(1)})^{-k_j'}$

(c") $\lim_{\epsilon \downarrow 0} \exp \limsup_{s\to\infty} \frac{1}{s} \log \mathbb{P}\left\{S_1'(\epsilon, 0) = s\right\} \leq p_e$.

Let $\mathcal{F}_i$ be a σ-field generated by $\beta[0], \cdots, \beta[i]$ and $g_\epsilon : \mathbb{R}^+ \to \mathbb{R}^+$ be the function of Lemma 7.24. Denote a random variable $d(\epsilon, N)$ as following:

$$d(\epsilon, N) := \sup_{l \in \mathbb{Z}, l - S_1'(\epsilon,0) \geq g_\epsilon(S_1'(\epsilon,0))} \frac{1}{N} \sum_{n=l+1}^{l+N} \mathbf{1}\left\{\left|\det\left(\begin{bmatrix} \mathbf{C_1}\mathbf{A_1}^{-k_1'} \\ \vdots \\ \mathbf{C_1}\mathbf{A_1}^{-k_{m_1-1}'} \\ \mathbf{C_1}\mathbf{A_1}^{-n} \end{bmatrix}\right)\right| < \epsilon^2 (\lambda_{m_1}^{(1)})^{-n} \prod_{1 \leq j \leq m_1-1} (\lambda_j^{(1)})^{-k_j} | \mathcal{F}_{S_1'(\epsilon,0)}\right\}.$$

Since (b") implies $cof_{m_1,m_1}\left(\begin{bmatrix} \mathbf{C_1}\mathbf{A_1}^{-k_1'} \\ \vdots \\ \mathbf{C_1}\mathbf{A_1}^{-k_{m_1-1}'} \\ \mathbf{C_1}\mathbf{A_1}^{-n} \end{bmatrix}\right) \geq \epsilon \prod_{1 \leq j \leq m_1-1} (\lambda_i^{(1)})^{-k_j'}$, by Lemma 7.24 we have

$$\lim_{\epsilon \downarrow 0} \lim_{N \to \infty} \operatorname{ess\,sup} d(\epsilon, N) = 0.$$

Denote $S_1''(\epsilon,0) := S_1'(\epsilon,0) + g_\epsilon(S_1'(\epsilon,0))$. From (ii) of Lemma 7.24 we know $g_\epsilon(k) \lesssim 1 + \log(k+1)$ for all $\epsilon > 0$. Therefore, by Lemma 7.1 we have

$$\lim_{\epsilon\downarrow 0} \exp \limsup_{s\to\infty} \frac{1}{s} \log \mathbb{P}\{S_1''(\epsilon,0) = s\} \le p_e. \tag{7.77}$$

Denote a stopping time

$$S_1'''(\epsilon,0) := \inf\left\{ n > S_1''(\epsilon,0) : \beta[n] = 1 \text{ and } \left|\det\left(\begin{bmatrix} \mathbf{C_1 A_1}^{-k_1'} \\ \vdots \\ \mathbf{C_1 A_1}^{-k'_{m_1-1}} \\ \mathbf{C_1 A_1}^{-n} \end{bmatrix}\right)\right| \ge \epsilon^2 (\lambda_{m_1}^{(1)})^{-n} \prod_{1\le j\le m_1-1} (\lambda_j^{(1)})^{-k_j'} \right\}.$$

Since $\beta[n]$ is a Bernoulli process,

$$\mathbb{P}\{S_1'''(\epsilon,0) - S_1''(\epsilon,0) \ge N | \mathcal{F}_{S_1''(\epsilon,0)}\} \le p_e^{N(1-d(\epsilon,N))}.$$

Therefore,

$$\lim_{\epsilon\downarrow 0} \exp \limsup_{N\to 0} \operatorname{ess\,sup} \frac{1}{N} \log \mathbb{P}\{S_1'''(\epsilon,0) - S_1''(\epsilon,0) \ge N | \mathcal{F}_{S_1''(\epsilon,0)}\} \le \lim_{\epsilon\downarrow 0} \lim_{N\to\infty} \operatorname{ess\,sup} p_e^{1-d(\epsilon,N)} \le p_e$$

i.e.

$$\lim_{\epsilon\downarrow 0} \exp \limsup_{s\to 0} \operatorname{ess\,sup} \frac{1}{s} \log \mathbb{P}\{S_1'''(\epsilon,0) - S_1''(\epsilon,0) = s | \mathcal{F}_{S_1''(\epsilon,0)}\} \le p_e. \tag{7.78}$$

By applying Lemma 7.2 to (7.77) and (7.78), we can conclude that

$$\lim_{\epsilon\downarrow 0} \exp \limsup_{s\to\infty} \frac{1}{s} \log \mathbb{P}\{S_1'''(\epsilon,0) = s\} \le p_e.$$

Therefore, if we denote $S_1(\epsilon,0) := S_1'''(\epsilon^{\frac{1}{2}},0)$, $S_1(\epsilon,0)$ satisfies the conditions of Claim 7.6 when we fix $k=0$.

Here, we know $\beta[n]$ is a stationary process. Thus, to prove the claim for general $k \in \mathbb{Z}^+$, we can shift the time index by k. Then, we can find a stopping time $S_1(\epsilon,k)$ such that for all $\epsilon > 0$ and $k \in \mathbb{Z}^+$ there exist $k \le k_1' < k_2' < \cdots < k_{m_1}' \le S_1(\epsilon,k)$ satisfying the following conditions:

(a"') $\beta[k_j'] = 1$ for $1 \le j \le m_1$

(b"') $\left|\det\left(\begin{bmatrix} \mathbf{C_1 A_1}^{-(k_1'-k)} \\ \mathbf{C_1 A_1}^{-(k_2'-k)} \\ \vdots \\ \mathbf{C_1 A_1}^{-(k'_{m_i}-k)} \end{bmatrix}\right)\right| \ge \epsilon \prod_{1\le j\le m_1} (\lambda_j^{(1)})^{-(k_j'-k)}$

(c"') $\lim_{\epsilon\downarrow 0} \exp \limsup_{s\to\infty} \sup_{k\in\mathbb{Z}^+} \frac{1}{s} \log \mathbb{P}\left\{S_1(\epsilon,k) - k = s\right\} \le p_e$.

Here, we can notice that the condition (b"') is equivalent to

$$\left|\det\left(\begin{bmatrix} \mathbf{C_1 A_1}^{-k_1'} \\ \mathbf{C_1 A_1}^{-k_2'} \\ \vdots \\ \mathbf{C_1 A_1}^{-k'_{m_i}} \end{bmatrix}\right)\right| \cdot \left|\det\left(\begin{bmatrix} \mathbf{A_1}^k \end{bmatrix}\right)\right| \ge \epsilon \prod_{1\le j\le m_1} (\lambda_j^{(1)})^{-(k_j'-k)}$$

$$(\Leftrightarrow) \left|\det\left(\begin{bmatrix} \mathbf{C_1 A_1}^{-k_1'} \\ \mathbf{C_1 A_1}^{-k_2'} \\ \vdots \\ \mathbf{C_1 A_1}^{-k'_{m_i}} \end{bmatrix}\right)\right| \ge \left|\det\left(\begin{bmatrix} \mathbf{A_1}^k \end{bmatrix}\right)\right|^{-1} \cdot \epsilon \prod_{1\le j\le m_1} (\lambda_j^{(1)})^{-(k_j'-k)} \ge \epsilon \prod_{1\le j\le m_1} (\lambda_j^{(1)})^{-k_j'}$$

Therefore, Claim 7.6 is true.

• Proof of Claim 7.5: By recursive use of Claim 7.6, we can find stopping times $S_2(\epsilon,k),\cdots,S_r(\epsilon,k)$ such that for all $\epsilon>0$ and $2\leq i\leq r$ there exist $S_{i-1}(\epsilon,k)<k_{i,1}<k_{i,2}<\cdots<k_{i,m_i}\leq S_i(\epsilon,k)$ satisfying the following condition:

(a) $\beta[k_{i,j}]=1$ for $1\leq j\leq m_i$

(b) $\left|\det\left(\begin{bmatrix}\mathbf{C_i}\mathbf{A_i}^{-k_{i,1}}\\ \mathbf{C_i}\mathbf{A_i}^{-k_{i,2}}\\ \vdots\\ \mathbf{C_i}\mathbf{A_i}^{-k_{i,m_i}}\end{bmatrix}\right)\right|\geq\epsilon\prod_{1\leq j\leq m_i}(\lambda_j^{(i)})^{-k_{i,j}}$

(c) $\lim_{\epsilon\downarrow 0}\exp\limsup_{s\to\infty}\operatorname{ess\,sup}\frac{1}{s}\log\mathbb{P}\{S_i(\epsilon,k)-S_{i-1}(\epsilon,k)=s|\mathcal{F}_{S_{i-1}(\epsilon,k)}\}\leq p_e$.

Then, by Lemma 7.2

$$\lim_{\epsilon\downarrow 0}\exp\limsup_{s\to\infty}\sup_{k\in\mathbb{Z}^+}\frac{1}{s}\log\mathbb{P}\{S_r(\epsilon,k)-k=s\}\leq p_e.$$

Therefore, if we denote $S(\epsilon,k):=S_r(\epsilon,k)$, $S(\epsilon,k)$ satisfies all the conditions of Claim 7.5. Thus, Claim 7.5 is true and the lemma is also true. □

We prove some properties about matrices which will be helpful in the proof of Lemma 2.3.

Lemma 7.26. *Let* $\mathbf{A}$ *and* $\mathbf{A}'$ *be Jordan block matrices with eigenvalues* $\lambda,\alpha\lambda(\alpha\neq 0)$ *respectively and the same size* $m\in\mathbb{N}$, *i.e.* $\mathbf{A}=\begin{bmatrix}\lambda & 1 & \cdots & 0\\ 0 & \lambda & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda\end{bmatrix}$ *and* $\mathbf{A}'=\begin{bmatrix}\alpha\lambda & 1 & \cdots & 0\\ 0 & \alpha\lambda & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \alpha\lambda\end{bmatrix}$. *Then, for all* $n\in\mathbb{Z}$

$$\mathbf{A'}^n=\begin{bmatrix}\alpha^{-(m-1)} & 0 & \cdots & 0\\ 0 & \alpha^{-(m-2)} & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 1\end{bmatrix}\mathbf{A}^n\begin{bmatrix}\alpha^{n+(m-1)} & 0 & \cdots & 0\\ 0 & \alpha^{n+(m-2)} & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \alpha^n\end{bmatrix}.$$

Proof.

$$\mathbf{A}'^n = \begin{bmatrix} (\alpha\lambda)^n & \binom{n}{1}(\alpha\lambda)^{n-1} & \binom{n}{2}(\alpha\lambda)^{n-2} & \cdots & \binom{n}{m}(\alpha\lambda)^{n-(m-1)} \\ 0 & (\alpha\lambda)^n & \binom{n}{1}(\alpha\lambda)^{n-1} & \cdots & \binom{n}{m-1}(\alpha\lambda)^{n-(m-2)} \\ 0 & 0 & (\alpha\lambda)^n & \cdots & \binom{n}{m-2}(\alpha\lambda)^{n-(m-3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & (\alpha\lambda)^n \end{bmatrix}$$

$$= \begin{bmatrix} \alpha^{-(m-1)} & 0 & 0 & \cdots & 0 \\ 0 & \alpha^{-(m-2)} & 0 & \cdots & 0 \\ 0 & 0 & \alpha^{-(m-3)} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

$$\cdot \begin{bmatrix} \alpha^{n+m-1}\lambda^n & \binom{n}{1}\alpha^{n-1+m-1}\lambda^{n-1} & \binom{n}{2}\alpha^{n-2+m-1}\lambda^{n-2} & \cdots & \binom{n}{m}\alpha^{n-(m-1)+m-1}\lambda^{n-(m-1)} \\ 0 & \alpha^{n+m-2}\lambda^n & \binom{n}{1}\alpha^{n-1+m-2}\lambda^{n-1} & \cdots & \binom{n}{m-1}\alpha^{n-(m-2)+m-2}\lambda^{n-(m-2)} \\ 0 & 0 & \alpha^{n+m-3}\lambda^n & \cdots & \binom{n}{m-2}\alpha^{n-(m-3)+m-3}\lambda^{n-(m-3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \alpha^n\lambda^n \end{bmatrix}$$

$$= \begin{bmatrix} \alpha^{-(m-1)} & 0 & 0 & \cdots & 0 \\ 0 & \alpha^{-(m-2)} & 0 & \cdots & 0 \\ 0 & 0 & \alpha^{-(m-3)} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

$$\cdot \begin{bmatrix} \alpha^{n+m-1}\lambda^n & \binom{n}{1}\alpha^{n+m-2}\lambda^{n-1} & \binom{n}{2}\alpha^{n+m-3}\lambda^{n-2} & \cdots & \binom{n}{m}\alpha^n\lambda^{n-m} \\ 0 & \alpha^{n+m-2}\lambda^n & \binom{n}{1}\alpha^{n+m-3}\lambda^{n-1} & \cdots & \binom{n}{m-1}\alpha^n\lambda^{n-(m-1)} \\ 0 & 0 & \alpha^{n+m-3}\lambda^n & \cdots & \binom{n}{m-2}\alpha^n\lambda^{n-(m-2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \alpha^n\lambda^n \end{bmatrix}$$

$$= \begin{bmatrix} \alpha^{-(m-1)} & 0 & 0 & \cdots & 0 \\ 0 & \alpha^{-(m-2)} & 0 & \cdots & 0 \\ 0 & 0 & \alpha^{-(m-3)} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

$$\cdot \begin{bmatrix} \lambda^n & \binom{n}{1}\lambda^{n-1} & \binom{n}{2}\lambda^{n-2} & \cdots & \binom{n}{m}\lambda^{n-m} \\ 0 & \lambda^n & \binom{n}{1}\lambda^{n-1} & \cdots & \binom{n}{m-1}\lambda^{n-(m-1)} \\ 0 & 0 & \lambda^n & \cdots & \binom{n}{m-2}\lambda^{n-(m-2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda^n \end{bmatrix} \cdot \begin{bmatrix} \alpha^{n+(m-1)} & 0 & 0 & \cdots & 0 \\ 0 & \alpha^{n+(m-2)} & 0 & \cdots & 0 \\ 0 & 0 & \alpha^{n+(m-3)} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \alpha^n \end{bmatrix}$$

$$= \begin{bmatrix} \alpha^{-(m-1)} & 0 & \cdots & 0 \\ 0 & \alpha^{-(m-2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \mathbf{A}^n \begin{bmatrix} \alpha^{n+(m-1)} & 0 & \cdots & 0 \\ 0 & \alpha^{n+(m-2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha^n \end{bmatrix}$$

This direct computation finishes the proof. $\square$

Lemma 7.27. *Let* $\mathbf{A}$ *be a Jordan block with eigenvalue* λ *and dimension* $m \times m$*. Then, the Jordan decomposition of the matrix* $\mathbf{A}^k$ *for* $k \in \mathbb{N}$ *is* $\mathbf{U}\mathbf{\Lambda}\mathbf{U}^{-1}$ *where* $\mathbf{U}$ *is an invertible upper triangular matrix —so the diagonal elements of* $\mathbf{U}$ *are non-zero— and* $\mathbf{\Lambda}$ *is a Jordan block with eigenvalue* λ^k *and dimension* $m \times m$*.*

Proof. We can see that $\mathbf{A}^k$ is a upper triangular toeplitz matrix whose diagonal elements are λ^k. Thus, $\det(s\mathbf{I} - \mathbf{A}^k) = (s - \lambda^k)^m$ and all eigenvalues of $\mathbf{A}^k$ are λ^k. Moreover, the rank of $\mathbf{A}^k - \lambda^k\mathbf{I}$ is $m - 1$. Thus, $\mathbf{\Lambda}$ has to be a Jordan block matrix with eigenvalue λ^k and dimension $m \times m$.

Moreover, $Ker\left((\mathbf{A} - \lambda^k\mathbf{I})^p\right) \supseteq span\{\mathbf{e_1}, \mathbf{e_2}, \cdots, \mathbf{e_p}\}$. Therefore, ith column of $\mathbf{U}^{-1}$ has to belong to the vector space $\{\mathbf{e_1}, \cdots, \mathbf{e_i}\}$ and $\mathbf{U}^{-1}$ is upper diagonal matrix. Here, the existence of the Jordan form of arbitrary matrices guarantee the invertibility of $\mathbf{U}$. Therefore, $\mathbf{U}$ is also an upper triangular matrix and the invertibility condition of an upper triangular matrix is its diagonal elements are non-zero. $\square$

Lemma 7.28. *Let* $\mathbf{A}$ *be a Jordan block matrix with eigenvalue* $\lambda \in \mathbb{C}$ *and size* $m \in \mathbb{N}$*, i.e.* $\mathbf{A} = \begin{bmatrix} \lambda & 1 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda \end{bmatrix}$. $\mathbf{C}$ *and* $\mathbf{C}'$ *are* $1 \times m$ *matrices such that*

$$\mathbf{C} = \begin{bmatrix} c_1 & c_2 & \cdots & c_m \end{bmatrix}$$
$$\mathbf{C}' = \begin{bmatrix} c'_1 & c'_2 & \cdots & c'_m \end{bmatrix}$$

where $c_i, c'_i \in \mathbb{C}$ *and* $c_1 \neq 0$*.*

For all $k \in \mathbb{R}$ *and* $m \times 1$ *matrices* $\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$ *and* $\mathbf{X}' = \begin{bmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_m \end{bmatrix}$*, there exists* $\mathbf{T}$ *such that*

$$(i)\mathbf{T} \text{ is an upper triangular matrix.}$$
$$(ii)\mathbf{C}\mathbf{A}^k\mathbf{X} + \mathbf{C}'\mathbf{A}^k\mathbf{X}' = \mathbf{C}\mathbf{A}^k\left(\mathbf{X} + \mathbf{T}\mathbf{X}'\right)$$

Moreover, the diagonal elements of $\mathbf{T}$ *are* $\frac{c'_1}{c_1}$*.*

Proof. Similar to Lemma 7.14. $\square$

Now, we can prove Lemma 2.3.

Proof of Lemma 2.3. We will prove the lemma by an induction on m, the dimension of the system. Recall that here we are using the definitions of (2.35), (2.36) for the system matrices $\mathbf{A}$, $\mathbf{C}$, $\mathbf{A_i}$, $\mathbf{C_i}$, $\cdots$.

(i) When $m = 1$,

In this case, the lemma reduces to the scalar problem and is trivially true. Precisely, if we choose $S_1(\epsilon, k)$ as $\inf\{s \geq k : \beta[s] = 1\}$, we can check all the conditions of the lemma are satisfied.

(ii) Now, we will assume the lemma is true when the system dimension is $m-1$ as an induction hypothesis, and prove the lemma holds for the system with dimension m.

Let $\mathbf{x_{i,j}}$ be a $m_{i,j} \times 1$ column vector, and $\mathbf{x}$ be $\begin{bmatrix} \mathbf{x_{1,1}} \\ \mathbf{x_{1,2}} \\ \vdots \\ \mathbf{x}_{\mu,\nu_\mu} \end{bmatrix}$. Here, $\mathbf{x}$ can be thought as the state of the system, and $\mathbf{x_{i,j}}$ corresponds to the states associated with the Jordan block $\mathbf{A_{i,j}}$. Recall that $\mathbf{A_{1,1}}$ is the Jordan block with the largest eigenvalue and size. For a vector $\mathbf{v}$, we also define $(\mathbf{v})_n$ as the nth element of $\mathbf{v}$.

The purpose of this proof is following: By Lemma 7.25, we already know that the lemma holds for systems with scalar observations and without eigenvalue cycles. Therefore, we first reduce the system to one with scalar observations and without eigenvalue cycles. To reduce the system to the one without eigenvalue cycles, we will use down-sampling ideas (polyphase decomposition) from signal processing [75]. To reduce the system to the one with scalar observations, we will multiply a proper post-processing matrix which combines vector observations into scalar observations. Then, we estimate the $m_{1,1}$th element of $\mathbf{x_{1,1}}$, which associated with the largest eigenvalue. Then, we subtract the estimate from the system. The resulting system becomes an $(m-1)$-dimensional system, and by the induction hypothesis, we can estimate the remaining states. As we mentioned before, this idea is called successive decoding in information theory [21].

Let's start with the down-sampling and reduction to scalar observations.

• Down-sampling the System by p and Reduction to Scalar Observations: The main difficulty in estimating the $m_{1,1}$th element of $\mathbf{x_{1,1}}$ is the periodicity of the system. To handle this difficulty, we down-sample the system. Let $p = \prod_{1 \leq i \leq \mu} p_i$. Recall that in (2.36), p_i was the period of each eigenvalue cycle. We can see when the system is down sampled by p, the resulting system becomes aperiodic. Thus, we can reduce the original periodic system to p aperiodic systems.

We can further reduce vector observation systems to scalar observation systems. Thus, the system reduces to an aperiodic system with scalar observations, and by Lemma 7.25 we can estimate the $m_{1,1}$th element of $\mathbf{x_{1,1}}$.

Since we are using induction for the proof, we can focus on the first eigenvalue cycle of the system.

Let $T_1, \cdots, T_R$ be all the sets T such that $T := \{t_1, \cdots, t_{|T|}\} \subseteq \{0, 1, \cdots, p_1 - 1\}$ and

$$\begin{bmatrix} \mathbf{C_1}\mathbf{A_1}^{-t_1} \\ \mathbf{C_1}\mathbf{A_1}^{-t_2} \\ \vdots \\ \mathbf{C_1}\mathbf{A_1}^{-t_{|T|}} \end{bmatrix} \text{ is full rank.} \tag{7.79}$$

Here, the definition of $\mathbf{A_1}$ and $\mathbf{C_1}$ is given in (2.36) and $\begin{bmatrix} \mathbf{C_1} diag\{\alpha_{1,1}, \cdots, \alpha_{1,\nu_1}\}^{-t_1} \\ \mathbf{C_1} diag\{\alpha_{1,1}, \cdots, \alpha_{1,\nu_1}\}^{-t_2} \\ \vdots \\ \mathbf{C_1} diag\{\alpha_{1,1}, \cdots, \alpha_{1,\nu_1}\}^{-t_{|T|}} \end{bmatrix}$ is also full rank. The number of such sets, R, is finite since p_1 is finite.

Therefore, for each $T_r := \{t_{r,1}, \cdots, t_{r,|T_r|}\}$ $(1 \le r \le R)$, we can find a matrix $\mathbf{L_r}$ such that

$$\mathbf{L_r} \begin{bmatrix} \mathbf{C_1} diag\{\alpha_{1,1}, \cdots, \alpha_{1,\nu_1}\}^{-t_{r,1}} \\ \mathbf{C_1} diag\{\alpha_{1,1}, \cdots, \alpha_{1,\nu_1}\}^{-t_{r,2}} \\ \vdots \\ \mathbf{C_1} diag\{\alpha_{1,1}, \cdots, \alpha_{1,\nu_1}\}^{-t_{r,|T_i|}} \end{bmatrix} = \mathbf{I}.$$

Let $\begin{bmatrix} \mathbf{L_{t_{r,1},r}} & \mathbf{L_{t_{r,2},r}} & \cdots & \mathbf{L_{t_{r,|T_r|},r}} \end{bmatrix}$ be the first row of $\mathbf{L_r}$ where $\mathbf{L_{t,r}}$ are $1 \times l$ matrices. Then,

$$\begin{bmatrix} \mathbf{L_{t_{r,1},r}} & \mathbf{L_{t_{r,2},r}} & \cdots & \mathbf{L_{t_{r,|T_r|},r}} \end{bmatrix} \begin{bmatrix} \mathbf{C_1} diag\{\alpha_{1,1}, \cdots, \alpha_{1,\nu_1}\}^{-t_{r,1}} \\ \mathbf{C_1} diag\{\alpha_{1,1}, \cdots, \alpha_{1,\nu_1}\}^{-t_{r,2}} \\ \vdots \\ \mathbf{C_1} diag\{\alpha_{1,1}, \cdots, \alpha_{1,\nu_1}\}^{-t_{r,|T_r|}} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}. \tag{7.80}$$

When $q \in \{0, 1, \cdots, p_1 - 1\} \setminus \{t_{r,1}, \cdots, t_{r,|T_r|}\}$, we put $\mathbf{L_{q,r}} = \mathbf{0}$. We also extend this definition of $\mathbf{L_{q,r}}$ to all $q \in \{0, \cdots, p-1\}, r \in \{1, \cdots, R\}$ by putting $\mathbf{L_{q,r}} := \mathbf{L_{q(\mod p_1),r}}$ for $q \ge p_1$. Then, we can easily check that (7.80) still holds as long as $t_{r,i}$ remains the same in $\mod p_1$.

Claim 7.7. *For a given $q \in \{0, \cdots, p-1\}$ and $r \in \{1, \cdots, R\}$, let $\mathbf{L_{q,r}C_1}$ be not $\mathbf{0}$. Then, there exist $\bar{\mathbf{C}}_{\mathbf{q,r}}$, $\bar{\mathbf{A}}_{\mathbf{q,r}}$, $\bar{\mathbf{U}}_{\mathbf{q,r}}$, $\bar{\mathbf{x}}_{\mathbf{q,r}}$ that satisfy the following conditions:*

(i) $\bar{\mathbf{A}}_{\mathbf{q,r}}$ is a $\bar{m}_{q,r} \times \bar{m}_{q,r}$ square matrix given in a Jordan form. The eigenvalues of $\bar{\mathbf{A}}_{\mathbf{q,r}}$ belong to $\{\lambda_{1,1}^p, \lambda_{2,1}^p, \cdots, \lambda_{\mu,1}^p\}$, and no two different Jordan blocks have the same eigenvalue. Therefore, $\bar{\mathbf{A}}_{\mathbf{q,r}}$ has no eigenvalue cycles. Furthermore, the first Jordan block(left-top) of $\bar{\mathbf{A}}_{\mathbf{q,r}}$ is a $m_{1,1} \times m_{1,1}$ Jordan block associated with eigenvalue $\lambda_{1,1}^p$.

(ii) $\bar{\mathbf{C}}_{\mathbf{q,r}}$ is a $1 \times \bar{m}_{q,r}$ row vector and $(\bar{\mathbf{A}}_{\mathbf{q,r}}, \bar{\mathbf{C}}_{\mathbf{q,r}})$ is observable.

(iii) $\bar{\mathbf{U}}_{\mathbf{q,r}}$ is a $\bar{m}_{q,r} \times \bar{m}_{q,r}$ invertible upper triangular matrix.

(iv) $\bar{\mathbf{x}}_{\mathbf{q},\mathbf{r}}$ *is a* $\bar{m}_{q,r} \times 1$ *column vector. There exists a nonzero constant* $g_{q,r}$ *such that*

$$(\bar{\mathbf{x}}_{\mathbf{q},\mathbf{r}})_{m_{1,1}} = g_{q,r} \left(\mathbf{L}_{\mathbf{q},\mathbf{r}} \mathbf{C}_{\mathbf{1}} diag\{\alpha_{1,1}, \cdots, \alpha_{1,\nu_1}\}^{-(q+(m_{1,1}-1))} \right) \begin{bmatrix} (\mathbf{x}_{\mathbf{1},\mathbf{1}})_{m_{1,1}} \\ (\mathbf{x}'_{\mathbf{1},\mathbf{2}})_{m_{1,1}} \\ \vdots \\ (\mathbf{x}'_{\mathbf{1},\nu_{\mathbf{1}}})_{m_{1,1}} \end{bmatrix}.$$

where $(\mathbf{x}'_{\mathbf{1},\mathbf{i}})_{m_{1,1}} = (\mathbf{x}_{\mathbf{1},\mathbf{i}})_{m_{1,1}}$ *when the size of* $\mathbf{x}_{\mathbf{1},\mathbf{i}}$ *is greater or equal to* $m_{1,1}$*, and* $(\mathbf{x}'_{\mathbf{1},\mathbf{i}})_{m_{1,1}} = 0$ *otherwise.*
(v) For all $k \in \mathbb{Z}^+$, $\mathbf{L}_{\mathbf{q},\mathbf{r}} \mathbf{C} \mathbf{A}^{-(pk+q)} \mathbf{x} = \bar{\mathbf{C}}_{\mathbf{q},\mathbf{r}} \bar{\mathbf{A}}_{\mathbf{q},\mathbf{r}}^{-k} \bar{\mathbf{U}}_{\mathbf{q},\mathbf{r}} \bar{\mathbf{x}}_{\mathbf{q},\mathbf{r}}$.

This claim tells that by sub-sampling with rate p, we get systems without eigenvalue cycles. Moreover, by multiplying the proper row vector to observations, we can reduce the system to a scalar observation system while keeping required information to estimate $(\mathbf{x}_{\mathbf{1},\mathbf{1}})_{m_{1,1}}$. When $\mathbf{L}_{\mathbf{q},\mathbf{r}} \mathbf{C}_{\mathbf{1}}$ is $\mathbf{0}$, the observation is not useful in estimation $(\mathbf{x}_{\mathbf{1},\mathbf{1}})_{m_{1,1}}$. Thus, we can ignore it.

Proof. The proof of the claim consists of two parts, down-sampling and reduction to a scalar observation system.

(1) Down-sampling the System by p:

By the definition of $\mathbf{C}$, $\mathbf{A}$, $\mathbf{C}_{\mathbf{i},\mathbf{j}}$, $\mathbf{A}_{\mathbf{i},\mathbf{j}}$, for all $k \in \mathbb{Z}$, $q \in \{0, \cdots, p-1\}$ we have

$$\mathbf{C}\mathbf{A}^{-(pk+q)}\mathbf{x} = \mathbf{C}_{\mathbf{1},\mathbf{1}} \mathbf{A}_{\mathbf{1},\mathbf{1}}{}^{-(pk+q)} \mathbf{x}_{\mathbf{1},\mathbf{1}} + \mathbf{C}_{\mathbf{1},\mathbf{2}} \mathbf{A}_{\mathbf{1},\mathbf{2}}{}^{-(pk+q)} \mathbf{x}_{\mathbf{1},\mathbf{2}} + \cdots + \mathbf{C}_{\mu,\nu_\mu} \mathbf{A}_{\mu,\nu_\mu}{}^{-(pk+q)} \mathbf{x}_{\mu,\nu_\mu} \tag{7.81}$$

Since the dimensions of $\mathbf{x}_{\mathbf{i},\mathbf{1}}, \cdots, \mathbf{x}_{\mathbf{i},\nu_{\mathbf{i}}}$ may be different, we will make them equal by extending the dimensions to the maximum, i.e. $m_{i,1}$. For the extension, we will append zeros at the end of the matrices. Let $\mathbf{C}'_{\mathbf{i},\mathbf{j}}$ be a $l \times m_{i,1}$ matrix given as $\begin{bmatrix} \mathbf{C}_{\mathbf{i},\mathbf{j}} & \mathbf{0}_{l \times (m_{i,1}-m_{i,j})} \end{bmatrix}$, $\mathbf{A}'_{\mathbf{i},\mathbf{j}}$ be a $m_{i,1} \times m_{i,1}$ Jordan block matrix with eigenvalue $\lambda_{i,j}$, and $\mathbf{x}'_{\mathbf{i},\mathbf{j}}$ be a $m_{i,1} \times 1$ column vector given as $\begin{bmatrix} \mathbf{x}_{\mathbf{i},\mathbf{j}} \\ \mathbf{0}_{(m_{i,1}-m_{i,j}) \times 1} \end{bmatrix}$. Then, by the construction, we can see that $(\mathbf{x}'_{\mathbf{1},\mathbf{1}})_{m_{1,1}} = (\mathbf{x}_{\mathbf{1},\mathbf{1}})_{m_{1,1}}$, and if $m_{1,i}$ is greater or equal to $m_{1,1}$ $(\mathbf{x}'_{\mathbf{1},\mathbf{i}})_{m_{1,1}} = (\mathbf{x}_{\mathbf{1},\mathbf{i}})_{m_{1,1}}$ and otherwise $(\mathbf{x}'_{\mathbf{1},\mathbf{i}})_{m_{1,1}} = 0$. Therefore, $\mathbf{x}'_{\mathbf{i},\mathbf{j}}$ satisfies the condition (iv) of the claim. Furthermore, the first column of $\mathbf{C}'_{\mathbf{i},\mathbf{j}}$ is equal to the first column of $\mathbf{C}_{\mathbf{i},\mathbf{j}}$ by construction.

We also define $\alpha_{i,j}$ to be $\frac{\lambda_{i,j}}{\lambda_{i,1}}$. Recall that $\lambda_{i,j}$ was defined as the eigenvalue corresponding to $\mathbf{A}_{\mathbf{i},\mathbf{j}}$ in (2.35). Then, by the definitions $\alpha_{i,j}^{p_i} = 1$.

Then, (7.81) can be written as follows:

$$
\begin{aligned}
&\mathbf{C}\mathbf{A}^{-(pk+q)}\mathbf{x} = \mathbf{C}'_{\mathbf{1},\mathbf{1}}{\mathbf{A}'_{\mathbf{1},\mathbf{1}}}^{-(pk+q)}\mathbf{x}'_{\mathbf{1},\mathbf{1}} + \mathbf{C}'_{\mathbf{1},\mathbf{2}}{\mathbf{A}'_{\mathbf{1},\mathbf{2}}}^{-(pk+q)}\mathbf{x}'_{\mathbf{1},\mathbf{2}} + \cdots + \mathbf{C}'_{\mu,\nu_\mu}{\mathbf{A}'_{\mu,\nu_\mu}}^{-(pk+q)}\mathbf{x}'_{\mu,\nu_\mu} \\
&= \mathbf{C}'_{\mathbf{1},\mathbf{1}}{\mathbf{A}'_{\mathbf{1},\mathbf{1}}}^{-(pk+q)}\mathbf{x}'_{\mathbf{1},\mathbf{1}} \\
&+ \mathbf{C}'_{\mathbf{1},\mathbf{2}}\begin{bmatrix} \alpha_{1,2}^{-(m_{1,1}-1)} & 0 & \cdots & 0 \\ 0 & \alpha_{1,2}^{-(m_{1,1}-2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}{\mathbf{A}'_{\mathbf{1},\mathbf{1}}}^{-(pk+q)} \\
&\cdot\begin{bmatrix} \alpha_{1,2}^{-(pk+q)+(m_{1,1}-1)} & 0 & \cdots & 0 \\ 0 & \alpha_{1,2}^{-(pk+q)+(m_{1,1}-2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_{1,2}^{-(pk+q)} \end{bmatrix}\mathbf{x}'_{\mathbf{1},\mathbf{2}} + \cdots \\
&+ \mathbf{C}'_{\mu,\nu_\mu}\begin{bmatrix} \alpha_{\mu,\nu_\mu}^{-(m_{\mu,1}-1)} & 0 & \cdots & 0 \\ 0 & \alpha_{\mu,\nu_\mu}^{-(m_{\mu,1}-2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}{\mathbf{A}'_{\mu,\mathbf{1}}}^{-(pk+q)} \\
&\cdot\begin{bmatrix} \alpha_{\mu,\nu_\mu}^{-(pk+q)+(m_{\mu,1}-1)} & 0 & \cdots & 0 \\ 0 & \alpha_{\mu,\nu_\mu}^{-(pk+q)+(m_{\mu,1}-2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_{\mu,\nu_\mu}^{-(pk+q)} \end{bmatrix}\mathbf{x}'_{\mu,\nu_\mu} \qquad (7.82)
\end{aligned}
$$

$$
\begin{aligned}
&= \mathbf{C}'_{\mathbf{1},\mathbf{1}}{\mathbf{A}'_{\mathbf{1},\mathbf{1}}}^{-(pk+q)}\mathbf{x}'_{\mathbf{1},\mathbf{1}} \\
&+ \mathbf{C}'_{\mathbf{1},\mathbf{2}}\begin{bmatrix} \alpha_{1,2}^{-(m_{1,1}-1)} & 0 & \cdots & 0 \\ 0 & \alpha_{1,2}^{-(m_{1,1}-2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}{\mathbf{A}'_{\mathbf{1},\mathbf{1}}}^{-(pk+q)}\begin{bmatrix} \alpha_{1,2}^{-q+(m_{1,1}-1)} & 0 & \cdots & 0 \\ 0 & \alpha_{1,2}^{-q+(m_{1,1}-2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_{1,2}^{-q} \end{bmatrix}\mathbf{x}'_{\mathbf{1},\mathbf{2}} + \cdots \\
&+ \mathbf{C}'_{\mu,\nu_\mu}\begin{bmatrix} \alpha_{\mu,\nu_\mu}^{-(m_{\mu,1}-1)} & 0 & \cdots & 0 \\ 0 & \alpha_{\mu,\nu_\mu}^{-(m_{\mu,1}-2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}{\mathbf{A}'_{\mu,\mathbf{1}}}^{-(pk+q)}\begin{bmatrix} \alpha_{\mu,\nu_\mu}^{-q+(m_{\mu,1}-1)} & 0 & \cdots & 0 \\ 0 & \alpha_{\mu,\nu_\mu}^{-q+(m_{\mu,1}-2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_{\mu,\nu_\mu}^{-q} \end{bmatrix}\mathbf{x}'_{\mu,\nu_\mu}. \qquad (7.83)
\end{aligned}
$$

Here, (7.82) follows from Lemma 7.26 and (7.83) follows from $\alpha_{i,j}^{p} = \left(\alpha_{i,j}^{p_i}\right)^{\prod_{j\neq i} p_j} = 1$. Recall that $m_{i,j}$ was defined as the size of $\mathbf{A_{i,j}}$ in (2.35).

Define

$$\mathbf{C}''_{\mathbf{i},\mathbf{j}} := \mathbf{C}'_{\mathbf{i},\mathbf{j}} \begin{bmatrix} \alpha_{i,j}^{-(m_{i,1}-1)} & 0 & \cdots & 0 \\ 0 & \alpha_{i,j}^{-(m_{i,1}-2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}, \quad (7.84)$$

$$\mathbf{x}''_{\mathbf{i},\mathbf{j}} := \begin{bmatrix} \alpha_{i,j}^{-q+(m_{i,1}-1)} & 0 & \cdots & 0 \\ 0 & \alpha_{i,j}^{-q+(m_{i,1}-2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_{i,j}^{-q} \end{bmatrix} \mathbf{x}'_{\mathbf{i},\mathbf{j}}. \quad (7.85)$$

Here, we can notice that the first column of $\mathbf{C}''_{\mathbf{i},\mathbf{j}}$ is $\alpha_{i,j}^{-(m_{i,1}-1)}$ times the first column of $\mathbf{C}'_{\mathbf{i},\mathbf{j}}$. Here, we know the first column of $\mathbf{C}'_{\mathbf{i},\mathbf{j}}$ is equal to the first column of $\mathbf{C}_{\mathbf{i},\mathbf{j}}$. The last element of $\mathbf{x}''_{\mathbf{i},\mathbf{j}}$ is $\alpha_{i,j}^{-q}$ times the last element of $\mathbf{x}'_{\mathbf{i},\mathbf{j}}$. (7.83) can be written as

$$\mathbf{C}\mathbf{A}^{-(pk+q)}\mathbf{x} = \mathbf{C}''_{\mathbf{1},\mathbf{1}}{\mathbf{A}'_{\mathbf{1},\mathbf{1}}}^{-(pk+q)}\mathbf{x}''_{\mathbf{1},\mathbf{1}} + \mathbf{C}''_{\mathbf{1},\mathbf{2}}{\mathbf{A}'_{\mathbf{1},\mathbf{1}}}^{-(pk+q)}\mathbf{x}''_{\mathbf{1},\mathbf{2}} + \cdots + \mathbf{C}''_{\mu,\nu_\mu}{\mathbf{A}'_{\mu,\mathbf{1}}}^{-(pk+q)}\mathbf{x}''_{\mu,\nu_\mu} \quad (7.86)$$

We can see all $\mathbf{x}''_{\mathbf{i},\mathbf{1}}, \cdots, \mathbf{x}''_{\mathbf{i},\nu_\mathbf{i}}$ are multiplied by the same matrix $\mathbf{A}'_{\mathbf{1},\mathbf{1}}$. Eventually, we will merge $\mathbf{x}''_{\mathbf{i},\mathbf{1}}, \cdots, \mathbf{x}''_{\mathbf{i},\nu_\mathbf{i}}$ by taking linear combinations.

(2) Reduction to the scalar observation: Now, we reduce $\mathbf{C}''_{\mathbf{i},\mathbf{j}}$ to row vectors by multiplying $\mathbf{L}_{\mathbf{q},\mathbf{r}}$ to (7.86).

$$\mathbf{L}_{\mathbf{q},\mathbf{r}}\mathbf{C}\mathbf{A}^{-(pk+q)}\mathbf{x} = \mathbf{L}_{\mathbf{q},\mathbf{r}}\mathbf{C}''_{\mathbf{1},\mathbf{1}}{\mathbf{A}'_{\mathbf{1},\mathbf{1}}}^{-(pk+q)}\mathbf{x}''_{\mathbf{1},\mathbf{1}} + \mathbf{L}_{\mathbf{q},\mathbf{r}}\mathbf{C}''_{\mathbf{1},\mathbf{2}}{\mathbf{A}'_{\mathbf{1},\mathbf{1}}}^{-(pk+q)}\mathbf{x}''_{\mathbf{1},\mathbf{2}} + \cdots + \mathbf{L}_{\mathbf{q},\mathbf{r}}\mathbf{C}''_{\mu,\nu_\mu}{\mathbf{A}'_{\mu,\mathbf{1}}}^{-(pk+q)}\mathbf{x}''_{\mu,\nu_\mu}. \quad (7.87)$$

Here, the systems $(\mathbf{A}'_{\mathbf{i},\mathbf{1}}, \mathbf{L}_{\mathbf{q},\mathbf{r}}\mathbf{C}''_{\mathbf{i},\mathbf{1}}), \cdots, (\mathbf{A}'_{\mathbf{i},\mathbf{1}}, \mathbf{L}_{\mathbf{q},\mathbf{r}}\mathbf{C}''_{\mathbf{i},\nu_\mathbf{i}})$ have the same dimension, but none of them might be observable. Therefore, we will make at least one of the systems be observable by truncation. Since $\mathbf{A}'_{\mathbf{i},\mathbf{1}}$ is a Jordan block matrix and $\mathbf{L}_{\mathbf{q},\mathbf{r}}\mathbf{C}''_{\mathbf{i},\mathbf{j}}$ is a row vector, $(\mathbf{A}'_{\mathbf{i},\mathbf{1}}, \mathbf{L}_{\mathbf{q},\mathbf{r}}\mathbf{C}''_{\mathbf{i},\mathbf{j}})$ is observable if and only if the first element of $\mathbf{L}_{\mathbf{q},\mathbf{r}}\mathbf{C}''_{\mathbf{i},\mathbf{j}}$ is not zero. Let m'_i be the smallest number such that at least one of the m'_ith elements of $\mathbf{L}_{\mathbf{q},\mathbf{r}}\mathbf{C}''_{\mathbf{i},\mathbf{1}}, \cdots, \mathbf{L}_{\mathbf{q},\mathbf{r}}\mathbf{C}''_{\mathbf{i},\nu_\mathbf{i}}$ becomes nonzero, and let $\mathbf{L}_{\mathbf{q},\mathbf{r}}\mathbf{C}''_{\mathbf{i},\nu_\mathbf{i}^\star}$ be the vector that achieves the minimum.

Then, we will reduce the dimension of $(\mathbf{A}'_{\mathbf{i},\mathbf{1}}, \mathbf{L}_{\mathbf{q},\mathbf{r}}\mathbf{C}''_{\mathbf{i},\nu_\mathbf{i}})$ by truncating the first $(m'_i - 1)$ vectors. Define $\mathbf{C}'''_{\mathbf{i},\mathbf{j}}$ as the matrix obtained by truncating the first $(m'_i - 1)$ columns of $\mathbf{C}''_{\mathbf{i},\mathbf{j}}$, $\mathbf{A}''_{\mathbf{i},\mathbf{j}}$ as the matrix obtained by truncating the first $(m'_i - 1)$ rows and columns of $\mathbf{A}'_{\mathbf{i},\mathbf{j}}$, and $\mathbf{x}'''_{\mathbf{i},\mathbf{j}}$ as the column vector obtained by truncating the first $(m'_i - 1)$ elements of $\mathbf{x}''_{\mathbf{i},\mathbf{j}}$.

In the claim, we assumed that $\mathbf{L}_{\mathbf{q},\mathbf{r}}\mathbf{C}_\mathbf{1}$ is not $\mathbf{0}$. Recall that the elements of $\mathbf{L}_{\mathbf{q},\mathbf{r}}\mathbf{C}_\mathbf{1}$ correspond to the first elements of $\mathbf{L}_{\mathbf{q},\mathbf{r}}\mathbf{C}_{\mathbf{1},\mathbf{1}}, \cdots, \mathbf{L}_{\mathbf{q},\mathbf{r}}\mathbf{C}_{\mathbf{1},\nu_\mathbf{1}}$, which are again equal to the first elements of $\mathbf{L}_{\mathbf{q},\mathbf{r}}\mathbf{C}'_{\mathbf{1},\mathbf{1}}, \cdots, \mathbf{L}_{\mathbf{q},\mathbf{r}}\mathbf{C}'_{\mathbf{1},\nu_\mathbf{1}}$. Since the first column of $\mathbf{C}''_{\mathbf{i},\mathbf{j}}$ is the first column of $\mathbf{C}'_{\mathbf{i},\mathbf{j}}$ times $\alpha_{i,j}^{-(m_{i,1}-1)}$, at least one of the systems $(\mathbf{A}'_{\mathbf{1},\mathbf{1}}, \mathbf{L}_{\mathbf{q},\mathbf{r}}\mathbf{C}''_{\mathbf{1},\mathbf{1}}), \cdots, (\mathbf{A}'_{\mathbf{1},\mathbf{1}}, \mathbf{L}_{\mathbf{q},\mathbf{r}}\mathbf{C}''_{\mathbf{1},\nu_\mathbf{1}})$ has to be observable.

Therefore, we can see $m'_1 = 1$ and

$$\mathbf{C}'''_{\mathbf{1},\mathbf{i}} = \mathbf{C}''_{\mathbf{1},\mathbf{i}}, \mathbf{A}''_{\mathbf{1},\mathbf{i}} = \mathbf{A}'_{\mathbf{1},\mathbf{i}}, \mathbf{x}'''_{\mathbf{1},\mathbf{i}} = \mathbf{x}''_{\mathbf{1},\mathbf{i}}. \quad (7.88)$$

Now, (7.87) becomes

$$\mathbf{L_{q,r}CA^{-(pk+q)}x} = \mathbf{L_{q,r}C'''_{1,1}A''_{1,1}}^{-(pk+q)}\mathbf{x'''_{1,1}} + \mathbf{L_{q,r}C'''_{1,2}A''_{1,1}}^{-(pk+q)}\mathbf{x'''_{1,2}} + \cdots + \mathbf{L_{q,r}C'''_{\mu,\nu_\mu}A''_{\mu,1}}^{-(pk+q)}\mathbf{x'''_{\mu,\nu_\mu}}.$$

Let $c'''_{i,j,1}$ be the first element of $\mathbf{L_{q,r}C'''_{i,j}}$. By Lemma 7.28, we can find upper triangular matrices $\mathbf{T_{i,j}}$ such that their diagonal elements are $\frac{c'''_{i,j,1}}{c'''_{i,\nu_i^\star,1}}$ and

$$\begin{aligned}\mathbf{L_{q,r}CA^{-(pk+q)}x} = &\mathbf{L_{q,r}C'''_{1,\nu_1^\star}A''_{1,1}}^{-(pk+q)}\left(\mathbf{T_{1,1}x'''_{1,1}} + \mathbf{T_{1,2}x'''_{1,2}} + \cdots + \mathbf{T}_{1,\nu_1}\mathbf{x'''_{1,\nu_1}}\right) + \cdots \\ &+ \mathbf{L_{q,r}C'''_{\mu,\nu_\mu^\star}A''_{\mu,1}}^{-(pk+q)}\left(\mathbf{T}_{\mu,1}\mathbf{x'''_{\mu,1}} + \mathbf{T}_{\mu,2}\mathbf{x'''_{\mu,2}} + \cdots + \mathbf{T}_{\mu,\nu_\mu}\mathbf{x'''_{\mu,\nu_\mu}}\right)\end{aligned} \tag{7.89}$$

where $c'''_{i,\nu_i^\star,1}$ is guaranteed to be nonzero by the construction.

Define $\mathbf{x''''_i}$ as

$$\left(\mathbf{T_{i,1}x'''_{i,1}} + \mathbf{T_{i,2}x'''_{i,2}} + \cdots + \mathbf{T}_{\mathbf{i},\nu_\mathbf{i}}\mathbf{x'''_{i,\nu_i}}\right). \tag{7.90}$$

Here, $\mathbf{A''_{i,1}}^{-(pk+q)}$ is not in a Jordan block. However, since $\mathbf{A''_{i,1}}$ is a Jordan block, by Lemma 7.27 the Jordan decomposition of $\mathbf{A''_{i,1}}^p$ is $\mathbf{U_i\Lambda_iU_i}^{-1}$ where $\mathbf{\Lambda_i}$ is a Jordan block whose eigenvalue is the pth power of the eigenvalue of $\mathbf{A''_{i,1}}$ and $\mathbf{U_i}$ is an upper triangular matrix whose diagonal entries are non-zero. Thus, (7.89) can be written as

$$\mathbf{L_{q,r}CA^{-(pk+q)}x} = \mathbf{L_{q,r}C'''_{1,\nu_1^\star}U_1\Lambda_1}^{-k}\mathbf{U_1}^{-1}\mathbf{A''_{1,1}}^{-q}\mathbf{x''''_1} + \cdots + \mathbf{L_{q,r}C'''_{\mu,\nu_\mu^\star}U_\mu\Lambda_\mu}^{-k}\mathbf{U}_\mu^{-1}\mathbf{A''_{\mu,1}}^{-q}\mathbf{x''''_\mu}$$

$$= \begin{bmatrix}\mathbf{L_{q,r}C'''_{1,\nu_1^\star}U_1} & \mathbf{L_{q,r}C'''_{2,\nu_2^\star}U_2} & \cdots & \mathbf{L_{q,r}C'''_{\mu,\nu_\mu^\star}U_\mu}\end{bmatrix}\begin{bmatrix}\mathbf{\Lambda_1} & 0 & \cdots & 0 \\ 0 & \mathbf{\Lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{\Lambda}_\mu\end{bmatrix}^{-k}$$

$$\cdot\begin{bmatrix}\mathbf{U_1}^{-1}\mathbf{A''_{1,1}}^{-q} & 0 & \cdots & 0 \\ 0 & \mathbf{U_2}^{-1}\mathbf{A''_{2,1}}^{-q} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{U}_\mu^{-1}\mathbf{A''_{\mu,1}}^{-q}\end{bmatrix}\begin{bmatrix}\mathbf{x''''_1} \\ \mathbf{x''''_2} \\ \vdots \\ \mathbf{x''''_\mu}\end{bmatrix}.$$

Let's define $\mathbf{\bar{C}_{q,r}}$ as $\begin{bmatrix}\mathbf{L_{q,r}C'''_{1,\nu_1^\star}U_1} & \mathbf{L_{q,r}C'''_{2,\nu_2^\star}U_2} & \cdots & \mathbf{L_{q,r}C'''_{\mu,\nu_\mu^\star}U_\mu}\end{bmatrix}$, $\mathbf{\bar{A}_{q,r}}$ as $\begin{bmatrix}\mathbf{\Lambda_1} & 0 & \cdots & 0 \\ 0 & \mathbf{\Lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{\Lambda}_\mu\end{bmatrix}$, $\mathbf{\bar{U}_{q,r}}$ as $\begin{bmatrix}\mathbf{U_1}^{-1}\mathbf{A''_{1,1}}^{-q} & 0 & \cdots & 0 \\ 0 & \mathbf{U_2}^{-1}\mathbf{A''_{2,1}}^{-q} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{U}_\mu^{-1}\mathbf{A''_{\mu,1}}^{-q}\end{bmatrix}$, $\mathbf{\bar{x}_{q,r}}$ as $\begin{bmatrix}\mathbf{x''''_1} \\ \mathbf{x''''_2} \\ \vdots \\ \mathbf{x''''_\mu}\end{bmatrix}$ and $\bar{m}_{q,r}$ as the dimension of $\mathbf{\bar{A}_{q,r}}$.

Here, we can see that $\bar{\mathbf{A}}_{\mathbf{q},\mathbf{r}}$ has no eigenvalue cycles and satisfies the condition (i) of the claim. Furthermore, since $\mathbf{U_i}$ is an upper triangular matrix whose diagonal elements are non-zero, the first elements of $\mathbf{L_{q,r}}\mathbf{C}'''_{\mathbf{i},\nu_\mathbf{i}^\star}\mathbf{U_i}$ are still non-zeros. Thus, the system $(\mathbf{\Lambda_i}, \mathbf{L_{q,r}}\mathbf{C}'''_{\mathbf{i},\nu_\mathbf{i}^\star}\mathbf{U_i})$ is observable and $(\bar{\mathbf{A}}_{\mathbf{q},\mathbf{r}}, \bar{\mathbf{C}}_{\mathbf{q},\mathbf{r}})$ is also observable, which satisfies the condition (ii) of the claim. We also have

$$\mathbf{L_{q,r}}\mathbf{C}\mathbf{A}^{-(pk+q)}\mathbf{x} = \bar{\mathbf{C}}_{\mathbf{q},\mathbf{r}}\bar{\mathbf{A}}_{\mathbf{q},\mathbf{r}}^{-k}\bar{\mathbf{U}}_{\mathbf{q},\mathbf{r}}\bar{\mathbf{x}}_{\mathbf{q},\mathbf{r}}$$

which is the condition (v) of the claim.

Let $c_{1,j,1}$ be the first element of $\mathbf{L_{q,r}}\mathbf{C_{1,j}}$. Then, we have

$$\begin{aligned}
(\bar{\mathbf{x}}_{\mathbf{q},\mathbf{r}})_{m_{1,1}} = (\mathbf{x}''''_{\mathbf{1}})_{m_{1,1}} &= \left(\frac{c'''_{1,1,1}}{c'''_{1,\nu_1^\star,1}}(\mathbf{x}'''_{\mathbf{1,1}})_{m_{1,1}} + \cdots + \frac{c'''_{1,\nu_1,1}}{c'''_{1,\nu_1^\star,1}}(\mathbf{x}'''_{\mathbf{1},\nu_\mathbf{1}})_{m_{1,1}} \right) && (7.91)\\
&= \left(\frac{c'''_{1,1,1}}{c'''_{1,\nu_1^\star,1}}\alpha_{1,1}^{-q}(\mathbf{x}'_{\mathbf{1,1}})_{m_{1,1}} + \cdots + \frac{c'''_{1,\nu_1,1}}{c'''_{1,\nu_1^\star,1}}\alpha_{1,\nu_1}^{-q}(\mathbf{x}'_{\mathbf{1},\nu_\mathbf{1}})_{m_{1,1}} \right) && (7.92)\\
&= \frac{1}{c'''_{1,\nu_1^\star,1}}\left(c_{1,1,1}\alpha_{1,1}^{-q-(m_{1,1}-1)}(\mathbf{x}'_{\mathbf{1,1}})_{m_{1,1}} + \cdots + c_{1,\nu_1,1}\alpha_{1,\nu_1}^{-q-(m_{1,1}-1)}(\mathbf{x}'_{\mathbf{1},\nu_\mathbf{1}})_{m_{1,1}} \right) && (7.93)\\
&= \frac{1}{c'''_{1,\nu_1^\star,1}}\left(\mathbf{L_{q,r}}\mathbf{C_1}diag\{\alpha_{1,1}, \cdots, \alpha_{1,\nu_1}\}^{-(q+(m_{1,1}-1))} \right) \begin{bmatrix} (\mathbf{x}'_{\mathbf{1,1}})_{m_{1,1}} \\ \vdots \\ (\mathbf{x}'_{\mathbf{1},\nu_\mathbf{1}})_{m_{1,1}} \end{bmatrix}
\end{aligned}$$

(7.91) follows from (7.90). (7.92) follows from (7.85), (7.88). (7.93) follows from (7.84), (7.88) and that the first column of $\mathbf{C}'_{\mathbf{i,j}}$ is the same as the first column of $\mathbf{C_{i,j}}$ as we mentioned above. Furthermore, as we mentioned above, $(\mathbf{x}'_{\mathbf{1,1}})_{m_{1,1}} = (\mathbf{x_{1,1}})_{m_{1,1}}$. Therefore, the condition (iv) of the claim is also satisfied, and this finishes the proof. □

• Estimating $(\mathbf{x})_{m_{1,1}}$: Now, we have systems without eigenvalue cycles and with scalar observations. Thus, by applying Lemma 7.25, we will estimate the state $(\mathbf{x})_{m_{1,1}}$.

Claim 7.8. *We can find a polynomial $\bar{p}(k)$, $\bar{m} \in \mathbb{N}$ and a family of stopping time $\{\bar{S}(\epsilon, k) : k \in \mathbb{Z}^+, \epsilon > 0\}$ such that for all $\epsilon > 0$, $k \in \mathbb{Z}^+$ there exist $k \leq \bar{k}_1 < \bar{k}_2 < \cdots < \bar{k}_{\bar{m}} \leq \bar{S}(\epsilon, k)$ and $\bar{\mathbf{M}}$ satisfying:*

(i) $\beta[\bar{k}_i] = 1$ *for* $1 \leq i \leq \bar{m}$

(ii) $\bar{\mathbf{M}} \begin{bmatrix} \mathbf{C}\mathbf{A}^{-\bar{k}_1} \\ \vdots \\ \mathbf{C}\mathbf{A}^{-\bar{k}_{\bar{m}}} \end{bmatrix} \mathbf{x} = (\mathbf{x})_{m_{1,1}}$

(iii) $\left|\bar{\mathbf{M}}\right|_{max} \leq \frac{\bar{p}(\bar{S}(\epsilon,k))}{\epsilon}|\lambda_{1,1}|^{\bar{S}(\epsilon,k)}$

(iv) $\lim_{\epsilon \downarrow 0} \exp \limsup_{s \to \infty} \sup_{k \in \mathbb{Z}^+} \frac{1}{s} \log \mathbb{P}\{\bar{S}(\epsilon, k) - k = s\} \leq p_e^{\frac{l_1}{p_1}}$

This claim tells that there exists an estimator $\bar{\mathbf{M}}$ for $(\mathbf{x})_{m_{1,1}}$ which use observations at time $\bar{k}_1, \cdots, \bar{k}_{\bar{m}}$.

Proof. For each $q \in \{0, \cdots, p-1\}$, we have the down-sampled systems $(\bar{\mathbf{A}}_{\mathbf{q,1}}, \bar{\mathbf{C}}_{\mathbf{q,1}}), \cdots, (\bar{\mathbf{A}}_{\mathbf{q,R}}, \bar{\mathbf{C}}_{\mathbf{q,R}})$ such that all systems are observable, $\bar{\mathbf{A}}_{\mathbf{q,i}}$ have no eigenvalue cycles, and $\bar{\mathbf{C}}_{\mathbf{q,i}}$ are row vectors. By Lemma 7.25, we can find a polynomial $p_q(k)$ and a family of random variables $\{\bar{S}_q(\epsilon, k) : k \in \mathbb{Z}^+, \epsilon > 0\}$ such that for all $\epsilon > 0$, $k \in \mathbb{Z}^+$ and $1 \le i \le R$ there exist $\lceil \frac{k-q}{p} \rceil \le k_{i,1} < k_{i,2} < \cdots < k_{i,\bar{m}_{q,i}} \le \bar{S}_q(\epsilon, k)$ and $\mathbf{M_i}$ satisfying:

(i) $\beta[pk_{i,j} + q] = 1$ for $1 \le j \le \bar{m}_{q,i}$

(ii) $\mathbf{M_i} \begin{bmatrix} \bar{\mathbf{C}}_{\mathbf{q,i}} \bar{\mathbf{A}}_{\mathbf{q,i}}^{-k_{i,1}} \\ \bar{\mathbf{C}}_{\mathbf{q,i}} \bar{\mathbf{A}}_{\mathbf{q,i}}^{-k_{i,2}} \\ \vdots \\ \bar{\mathbf{C}}_{\mathbf{q,i}} \bar{\mathbf{A}}_{\mathbf{q,i}}^{-k_{i,\bar{m}_{q,i}}} \end{bmatrix} = \mathbf{I}_{\bar{\mathbf{m}}_{\mathbf{q,i}} \times \bar{\mathbf{m}}_{\mathbf{q,i}}}$

(iii) $|\mathbf{M_i}|_{max} \le \frac{p_q(\bar{S}_q(\epsilon,k))}{\epsilon} (|\lambda_{1,1}|^p)^{\bar{S}_q(\epsilon,k)}$

(iv) $\lim_{\epsilon \downarrow 0} \exp \limsup_{s \to \infty} \sup_{k \in \mathbb{Z}^+} \frac{1}{s} \log \mathbb{P}\{\bar{S}_q(\epsilon, k) - \lceil \frac{k-q}{p} \rceil = s\} = p_e$.

By the property (iv) of $\bar{S}_q(\epsilon, k)$, we get

$$\lim_{\epsilon \downarrow 0} \exp \limsup_{s \to \infty} \sup_{k \in \mathbb{Z}^+} \frac{1}{s} \log \mathbb{P}\{p\bar{S}_q(\epsilon, k) - p\lceil \frac{k-q}{p} \rceil = s\} = p_e^{\frac{1}{p}}$$

which implies

$$\lim_{\epsilon \downarrow 0} \exp \limsup_{s \to \infty} \sup_{k \in \mathbb{Z}^+} \frac{1}{s} \log \mathbb{P}\{(p\bar{S}_q(\epsilon, k) + q) - k = s\} = p_e^{\frac{1}{p}}.$$

Moreover, $\bar{S}_q(\epsilon, k)$ depends on only $\beta[q], \beta[p+q], \beta[2p+q], \cdots$. Thus, $\bar{S}_0(\epsilon, k), \cdots, \bar{S}_{p-1}(\epsilon, k)$ are independent.

Now, we can estimate the state of each sub-sampled system. We will leverage these estimations to the estimation of the state $(\mathbf{x})_{m_{1,1}}$.

First, notice that the down-sampling rate p is much larger than p_1. Therefore, we make the corresponding definition to (7.79) for the longer period p. Let $T'_1, \cdots, T'_{R'}$ be all the sets T' such that $T' := \{t'_1, \cdots, t'_{|T'|}\} \subseteq \{0, 1, \cdots, p-1\}$ and

$$\begin{bmatrix} \mathbf{C_1} \mathbf{A_1}^{-t'_1} \\ \mathbf{C_1} \mathbf{A_1}^{-t'_2} \\ \vdots \\ \mathbf{C_1} \mathbf{A_1}^{-t'_{|T'|}} \end{bmatrix} \text{ is full rank.}$$

Here, we can ask how many observations have to be erased to make the observability Gramian of $(\mathbf{A_1}, \mathbf{C_1})$ rank deficient during the period p. Obviously, the answer is $l_1 \prod_{2 \le j \le \mu} p_j$ where the definition of l_1 is shown in (2.37). The reason for this is that we have to erase at least l_1 observations for each period p_1 to make the observability Gramian rank deficient. Formally, it can be written as follows:

$$min\{|T| : T = \{t_1, \cdots, t_{|T|}\} \subseteq \{0, 1, \cdots, p-1\}, T'_i \not\subseteq T \text{ for all } 1 \le i \le R'\} = l_1 \prod_{2 \le j \le \mu} p_j.$$

Denote a stopping time $\bar{S}(\epsilon,k)$ as the minimum time until we have enough observations to make the observability Gramian of $(\mathbf{A_1},\mathbf{C_1})$ full rank. Formally,

$$\bar{S}(\epsilon,k)-k := \inf\{s : \exists i \in \{1,\cdots,R'\} \text{ s.t. } T_i' = \{t_1',t_2',\cdots t_{|T_i'|}'\} \text{ and}$$
$$(p\bar{S}_{t_1'}(\epsilon,k)+t_1')-k \le s, (p\bar{S}_{t_2'}(\epsilon,k)+t_2')-k \le s,\cdots,(p\bar{S}_{t_{|T_i'|}'}(\epsilon,k)+t_{|T_i'|}')-k \le s\}.$$

Then, by Lemma 7.3 we have

$$\lim_{\epsilon\downarrow 0}\exp\limsup_{s\to\infty}\sup_{k\in\mathbb{Z}^+}\frac{1}{s}\log\mathbb{P}\{\bar{S}(\epsilon,k)-k=s\} \le p_e^{\frac{l_1\prod_{j\neq 1}p_j}{p}} = p_e^{\frac{l_1}{p_1}}.$$

Without loss of generality, let T_1' be the set that satisfies the definition of $\bar{S}(\epsilon,k)$. Then, by the definition of T_1' and T_i, there must exist T_i such that T_1' contains T_i in mod p_1. Let T_1 be such a set without loss of generality. Then, we can find $\{t_1',\cdots,t_{|T_1|}'\}$ which is included in T_1' and includes T_1 in mod p_1. Formally, $\{t_1',\cdots,t_{|T_1|}'\}\subseteq T_1'$ and $\{t_1'(mod\ p_1),\cdots,t_{|T_1|}'(mod\ p_1)\}=T_1$.

Then, from the definition of $\bar{S}(\epsilon,k)$ and $\bar{S}_q(\epsilon,k)$, for each $q\in\{t_1',\cdots,t_{|T_1|}'\}$ we can find $\lceil\frac{k-q}{p}\rceil \le k_{q,1} < k_{q,2} < \cdots < k_{q,\bar{m}_{q,1}} \le \bar{S}_q(\epsilon,k)$ and $\mathbf{M_q}$ satisfying the following conditions:

(i') $\beta[pk_{q,j}+q]=1$ for $1\le j\le \bar{m}_{q,1}$

(ii') $\mathbf{M_q}\begin{bmatrix}\mathbf{\bar{C}_{q,1}}\mathbf{\bar{A}_{q,1}}^{-k_{q,1}} \\ \mathbf{\bar{C}_{q,1}}\mathbf{\bar{A}_{q,1}}^{-k_{q,2}} \\ \vdots \\ \mathbf{\bar{C}_{q,1}}\mathbf{\bar{A}_{q,1}}^{-k_{q,\bar{m}_{q,1}}}\end{bmatrix} = \mathbf{I_{\bar{m}_{q,1}\times\bar{m}_{q,1}}}$

(iii') $|\mathbf{M_q}|_{max} \le \frac{p_q(\bar{S}_q(\epsilon,k))}{\epsilon}(|\lambda_{1,1}|^p)^{\bar{S}_q(\epsilon,k)}$.

(iv') $p\bar{S}_q(\epsilon,k)+q \le \bar{S}(\epsilon,k)$

Then, we have

$$
\begin{aligned}
&diag\{\bar{\mathbf{U}}^{-1}_{\mathbf{t}'_1,\mathbf{1}}\mathbf{M}_{\mathbf{t}'_1},\bar{\mathbf{U}}^{-1}_{\mathbf{t}'_2,\mathbf{1}}\mathbf{M}_{\mathbf{t}'_2},\cdots,\bar{\mathbf{U}}^{-1}_{\mathbf{t}'_{|\mathbf{T_1}|},\mathbf{1}}\mathbf{M}_{\mathbf{t}'_{|\mathbf{T_1}|}}\}diag\{\mathbf{L}_{\mathbf{t}'_1,\mathbf{1}},\mathbf{L}_{\mathbf{t}'_1,\mathbf{1}},\cdots,\mathbf{L}_{\mathbf{t}'_{|\mathbf{T_1}|},\mathbf{1}}\} \\
&\cdot\begin{bmatrix}
\mathbf{CA}^{-(pk_{t'_1,1}+t'_1)} \\
\mathbf{CA}^{-(pk_{t'_1,2}+t'_1)} \\
\vdots \\
\mathbf{CA}^{-(pk_{t'_1,\bar{m}_{t'_1,1}}+t'_1)} \\
\mathbf{CA}^{-(pk_{t'_2,1}+t'_2)} \\
\vdots \\
\mathbf{CA}^{-(pk_{t'_{|T_1|},\bar{m}_{t'_{|T_1|},1}}+t'_{|T_1|})}
\end{bmatrix}\mathbf{x} \\
&= diag\{\bar{\mathbf{U}}^{-1}_{\mathbf{t}'_1,\mathbf{1}}\mathbf{M}_{\mathbf{t}'_1},\bar{\mathbf{U}}^{-1}_{\mathbf{t}'_2,\mathbf{1}}\mathbf{M}_{\mathbf{t}'_2},\cdots,\bar{\mathbf{U}}^{-1}_{\mathbf{t}'_{|\mathbf{T_1}|},\mathbf{1}}\mathbf{M}_{\mathbf{t}'_{|\mathbf{T_1}|}}\}
\begin{bmatrix}
\mathbf{L}_{\mathbf{t}'_1,\mathbf{1}}\mathbf{CA}^{-(pk_{t'_1,1}+t'_1)}\mathbf{x} \\
\mathbf{L}_{\mathbf{t}'_1,\mathbf{1}}\mathbf{CA}^{-(pk_{t'_1,2}+t'_1)}\mathbf{x} \\
\vdots \\
\mathbf{L}_{\mathbf{t}'_1,\mathbf{1}}\mathbf{CA}^{-(pk_{t'_1,\bar{m}_{t'_1,1}}+t'_1)}\mathbf{x} \\
\mathbf{L}_{\mathbf{t}'_2,\mathbf{1}}\mathbf{CA}^{-(pk_{t'_2,1}+t'_2)}\mathbf{x} \\
\vdots \\
\mathbf{L}_{\mathbf{t}'_{|\mathbf{T_1}|},\mathbf{1}}\mathbf{CA}^{-(pk_{t'_{|T_1|},\bar{m}_{t'_{|T_1|},1}}+t'_{|T_1|})}\mathbf{x}
\end{bmatrix} \\
&= diag\{\bar{\mathbf{U}}^{-1}_{\mathbf{t}'_1,\mathbf{1}}\mathbf{M}_{\mathbf{t}'_1},\bar{\mathbf{U}}^{-1}_{\mathbf{t}'_2,\mathbf{1}}\mathbf{M}_{\mathbf{t}'_2},\cdots,\bar{\mathbf{U}}^{-1}_{\mathbf{t}'_{|\mathbf{T_1}|},\mathbf{1}}\mathbf{M}_{\mathbf{t}'_{|\mathbf{T_1}|}}\}
\begin{bmatrix}
\bar{\mathbf{C}}_{\mathbf{t}'_1,\mathbf{1}}\bar{\mathbf{A}}^{-k_{t'_1,1}}_{\mathbf{t}'_1,\mathbf{1}}\bar{\mathbf{U}}_{\mathbf{t}'_1,\mathbf{1}}\bar{\mathbf{x}}_{\mathbf{t}'_1,\mathbf{1}} \\
\bar{\mathbf{C}}_{\mathbf{t}'_1,\mathbf{1}}\bar{\mathbf{A}}^{-k_{t'_1,2}}_{\mathbf{t}'_1,\mathbf{1}}\bar{\mathbf{U}}_{\mathbf{t}'_1,\mathbf{1}}\bar{\mathbf{x}}_{\mathbf{t}'_1,\mathbf{1}} \\
\vdots \\
\bar{\mathbf{C}}_{\mathbf{t}'_1,\mathbf{1}}\bar{\mathbf{A}}^{-k_{t'_1,\bar{m}_{t'_1,1}}}_{\mathbf{t}'_1,\mathbf{1}}\bar{\mathbf{U}}_{\mathbf{t}'_1,\mathbf{1}}\bar{\mathbf{x}}_{\mathbf{t}'_1,\mathbf{1}} \\
\bar{\mathbf{C}}_{\mathbf{t}'_2,\mathbf{1}}\bar{\mathbf{A}}^{-k_{t'_2,1}}_{\mathbf{t}'_2,\mathbf{1}}\bar{\mathbf{U}}_{\mathbf{t}'_2,\mathbf{1}}\bar{\mathbf{x}}_{\mathbf{t}'_2,\mathbf{1}} \\
\vdots \\
\bar{\mathbf{C}}_{\mathbf{t}'_{|\mathbf{T_1}|},\mathbf{1}}\bar{\mathbf{A}}^{-k_{t'_{|T_1|},\bar{m}_{t'_{|T_1|},1}}}_{\mathbf{t}'_{|\mathbf{T_1}|},\mathbf{1}}\bar{\mathbf{U}}_{\mathbf{t}'_{|\mathbf{T_1}|},\mathbf{1}}\bar{\mathbf{x}}_{\mathbf{t}'_{|\mathbf{T_1}|},\mathbf{1}}
\end{bmatrix}
\end{aligned}
\tag{7.94}
$$

$$
\begin{aligned}
&= \begin{bmatrix}
\bar{\mathbf{U}}^{-1}_{\mathbf{t}'_1,\mathbf{1}}\mathbf{M}_{\mathbf{t}'_1}
\begin{bmatrix}
\bar{\mathbf{C}}_{\mathbf{t}'_1,\mathbf{1}}\bar{\mathbf{A}}^{-k_{t'_1,1}}_{\mathbf{t}'_1,\mathbf{1}} \\
\vdots \\
\bar{\mathbf{C}}_{\mathbf{t}'_1,\mathbf{1}}\bar{\mathbf{A}}^{-k_{t'_1,\bar{m}_{t'_1,1}}}_{\mathbf{t}'_1,\mathbf{1}}
\end{bmatrix}\bar{\mathbf{U}}_{\mathbf{t}'_1,\mathbf{1}}\bar{\mathbf{x}}_{\mathbf{t}'_1,\mathbf{1}} \\
\vdots \\
\bar{\mathbf{U}}^{-1}_{\mathbf{t}'_{|\mathbf{T_1}|},\mathbf{1}}\mathbf{M}_{\mathbf{t}'_{|\mathbf{T_1}|}}
\begin{bmatrix}
\bar{\mathbf{C}}_{\mathbf{t}'_{|\mathbf{T_1}|},\mathbf{1}}\bar{\mathbf{A}}^{-k_{t'_{|T_1|},1}}_{\mathbf{t}'_1,\mathbf{1}} \\
\vdots \\
\bar{\mathbf{C}}_{\mathbf{t}'_{|\mathbf{T_1}|},\mathbf{1}}\bar{\mathbf{A}}^{-k_{t'_{|T_1|},\bar{m}_{t'_{|T_1|},1}}}_{\mathbf{t}'_1,\mathbf{1}}
\end{bmatrix}\bar{\mathbf{U}}_{\mathbf{t}'_{|\mathbf{T_1}|},\mathbf{1}}\bar{\mathbf{x}}_{\mathbf{t}'_{|\mathbf{T_1}|},\mathbf{1}}
\end{bmatrix} \\
&= \begin{bmatrix}
\bar{\mathbf{x}}_{\mathbf{t}'_1,\mathbf{1}} \\
\vdots \\
\bar{\mathbf{x}}_{\mathbf{t}'_{|\mathbf{T_1}|},\mathbf{1}}
\end{bmatrix}.
\end{aligned}
\tag{7.95}
$$

Here, (7.94) comes from the condition (v) of Claim 7.7. (7.95) comes from the definition of $\mathbf{M_q}$.

Now, we will estimate $(\mathbf{x})_{m_{1,1}}$ based on $\bar{\mathbf{x}}_{\mathbf{t'_1},\mathbf{1}}, \cdots, \bar{\mathbf{x}}_{\mathbf{t'_{|T_1|}},\mathbf{1}}$. Let $\mathbf{e}_{\mathbf{m_{1,1}}}^{\bar{\mathbf{m}}_{\mathbf{q},\mathbf{r}}}$ be a $1 \times \bar{m}_{q,r}$ row vector whose elements are all zeros except $m_{1,1}$th element which is 1. Then, we have the following equation:

$$
\begin{aligned}
&\begin{bmatrix} \frac{1}{g_{t'_1,1}}\mathbf{e}_{\mathbf{m_{1,1}}}^{\bar{\mathbf{m}}_{\mathbf{t'_1},\mathbf{1}}} & \cdots & \frac{1}{g_{t'_{|T_1|},1}}\mathbf{e}_{\mathbf{m_{1,1}}}^{\bar{\mathbf{m}}_{|\mathbf{t'_{|T_1|}}|,\mathbf{1}}} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{x}}_{\mathbf{t'_1},\mathbf{1}} \\ \vdots \\ \bar{\mathbf{x}}_{\mathbf{t'_{|T_1|}},\mathbf{1}} \end{bmatrix} \\
&= \frac{1}{g_{t'_1,1}}(\bar{\mathbf{x}}_{\mathbf{t'_1},\mathbf{1}})_{m_{1,1}} + \cdots + \frac{1}{g_{t'_{|T_1|},1}}(\bar{\mathbf{x}}_{\mathbf{t'_{|T_1|}},\mathbf{1}})_{m_{1,1}} \\
&= \left(\mathbf{L}_{\mathbf{t'_1},\mathbf{1}}\mathbf{C_1} diag\{\alpha_{1,1},\cdots,\alpha_{1,\nu_1}\}^{-(t'_1+(m_{1,1}-1))}\right) \begin{bmatrix} (\mathbf{x_{1,1}})_{m_{1,1}} \\ \vdots \\ (\mathbf{x'_{1,\nu_1}})_{m_{1,1}} \end{bmatrix} + \cdots \\
&+ \left(\mathbf{L}_{\mathbf{t'_{|T_1|}},\mathbf{1}}\mathbf{C_1} diag\{\alpha_{1,1},\cdots,\alpha_{1,\nu_1}\}^{-(t'_{|T_1|}+(m_{1,1}-1))}\right) \begin{bmatrix} (\mathbf{x_{1,1}})_{m_{1,1}} \\ \vdots \\ (\mathbf{x'_{1,\nu_1}})_{m_{1,1}} \end{bmatrix} \qquad (7.96) \\
&= \begin{bmatrix} \mathbf{L}_{\mathbf{t'_1},\mathbf{1}} & \cdots & \mathbf{L}_{\mathbf{t'_{|T_1|}},\mathbf{1}} \end{bmatrix} \begin{bmatrix} \mathbf{C_1} diag\{\alpha_{1,1},\cdots,\alpha_{1,\nu_1}\}^{-t'_1} \\ \vdots \\ \mathbf{C_1} diag\{\alpha_{1,1},\cdots,\alpha_{1,\nu_1}\}^{-t'_{|T_1|}} \end{bmatrix} \begin{bmatrix} \alpha_{1,1}^{-m_{1,1}+1}(\mathbf{x_{1,1}})_{m_{1,1}} \\ \vdots \\ \alpha_{1,\nu_1}^{-m_{1,1}+1}(\mathbf{x'_{1,\nu_1}})_{m_{1,1}} \end{bmatrix} \\
&= \alpha_{1,1}^{-m_{1,1}+1}(\mathbf{x_{1,1}})_{m_{1,1}} = \alpha_{1,1}^{-m_{1,1}+1}(\mathbf{x})_{m_{1,1}}. \qquad (7.97)
\end{aligned}
$$

Here, (7.96) follows from the condition (iv) of Claim 7.7. (7.97) follows from (7.80) and $\{t'_1(mod\ p_1), \cdots, t'_{|T_1|}(mod\ p_1)\} = T_1$.

Now, we merge the results from (7.95) and (7.97) to make an estimator for $(\mathbf{x})_{m_{1,1}}$. Define

$$
\begin{aligned}
\bar{\mathbf{M}} :=& \alpha_{1,1}^{m_{1,1}-1} \begin{bmatrix} \frac{1}{g_{t'_1,1}}\mathbf{e}_{\mathbf{m_{1,1}}}^{\bar{\mathbf{m}}_{|\mathbf{t'_1}|,\mathbf{1}}} & \cdots & \frac{1}{g_{t'_{|T_1|},1}}\mathbf{e}_{\mathbf{m_{1,1}}}^{\bar{\mathbf{m}}_{|\mathbf{t'_{|T_1|}}|,\mathbf{1}}} \end{bmatrix} \\
&\cdot diag\{\bar{\mathbf{U}}_{\mathbf{t'_1},\mathbf{1}}^{-1}\mathbf{M}_{\mathbf{t'_1}}, \bar{\mathbf{U}}_{\mathbf{t'_2},\mathbf{1}}^{-1}\mathbf{M}_{\mathbf{t'_2}}, \cdots, \bar{\mathbf{U}}_{\mathbf{t'_{|T_1|}},\mathbf{1}}^{-1}\mathbf{M}_{\mathbf{t'_{|T_1|}}}\} diag\{\mathbf{L}_{\mathbf{t'_1},\mathbf{1}}, \mathbf{L}_{\mathbf{t'_1},\mathbf{1}}, \cdots, \mathbf{L}_{\mathbf{t'_{|T_1|}},\mathbf{1}}\}
\end{aligned}
$$

and

$$
\begin{bmatrix} \mathbf{CA}^{-\bar{k}_1} \\ \vdots \\ \mathbf{CA}^{-\bar{k}_{\bar{m}}} \end{bmatrix} := \begin{bmatrix} \mathbf{CA}^{-(pk_{t'_1,1}+t'_1)} \\ \mathbf{CA}^{-(pk_{t'_1,2}+t'_1)} \\ \vdots \\ \mathbf{CA}^{-(pk_{t'_1,\bar{m}_{t'_1,1}}+t'_1)} \\ \mathbf{CA}^{-(pk_{t'_2,1}+t'_2)} \\ \vdots \\ \mathbf{CA}^{-(pk_{t'_{|T_1|},\bar{m}_{t'_{|T_1|},1}}+t'_{|T_1|})} \end{bmatrix}.
$$

Then, by (iii') and (iv') we can find a positive polynomial $\bar{p}(k)$ such that

$$|\bar{\mathbf{M}}|_{max} \lesssim \max_{1 \leq i \leq |T_1|} \{|\mathbf{M}_{\mathbf{t}'_i}|_{max}\} \leq \frac{\bar{p}(\bar{S}(\epsilon,k))}{\epsilon} |\lambda_{1,1}|^{\bar{S}(\epsilon,k)}.$$

Moreover, by (7.95) and (7.97) we have

$$\bar{\mathbf{M}} \begin{bmatrix} \mathbf{C}\mathbf{A}^{-\bar{k}_1} \\ \vdots \\ \mathbf{C}\mathbf{A}^{-\bar{k}_{\bar{m}}} \end{bmatrix} \mathbf{x} = (\mathbf{x})_{m_{1,1}}.$$

This finishes the proof of the claim $\square$

• Subtracting $(\mathbf{x})_{m_{1,1}}$ from the observations: Now, we have an estimation for $(\mathbf{x})_{m_{1,1}}$. We will remove it from the system.

$\widetilde{\mathbf{A}}$, $\widetilde{\mathbf{C}}$ and $\widetilde{\mathbf{x}}$ are the system matrices after the removal. Formally, $\widetilde{\mathbf{A}}$ is obtained by removing the $m_{1,1}$th row and column from $\mathbf{A}$, $\widetilde{\mathbf{C}}$ is obtained by removing the $m_{1,1}$th row from $\mathbf{C}$, and $\widetilde{\mathbf{x}}$ is obtained by removing the $m_{1,1}$th component from $\mathbf{x}$ respectively.

Denote the $m_{1,1}$th column of $\mathbf{C}\mathbf{A}^{-k}$ as $\mathbf{R}(k)$. Then, we have the following relation between the original system $(\mathbf{A}, \mathbf{C})$ and the new system $(\widetilde{\mathbf{A}}, \widetilde{\mathbf{C}})$:

$$\mathbf{C}\mathbf{A}^{-k}\mathbf{x} - \mathbf{R}(k)(\mathbf{x})_{m_{1,1}} = \widetilde{\mathbf{C}}\widetilde{\mathbf{A}}^{-k}\widetilde{\mathbf{x}} \tag{7.98}$$

which can be easily proved from the block diagonal structure of $\mathbf{A}$. From the definition of $\mathbf{R}(k)$, we can further see that there exists a polynomial $\widetilde{p}(k)$ such that $|\mathbf{R}(k)|_{max} \leq \widetilde{p}(k)|\lambda_{1,1}|^{-k}$. Thus, when $|\lambda_{1,1}| > 1$ we can find a threshold $k_{th} \geq 0$ such that all $k \geq k_{th}$, $\widetilde{p}(k)|\lambda_{1,1}|^{-k}$ is a decreasing function. When $|\lambda_{1,1}| = 1$, we simply put $k_{th} = 0$.

• Decoding the remaining element of $\mathbf{x}$: We decoded and subtracted the state $(\mathbf{x})_{m_{1,1}}$ from the system. After subtracting, the remaining system matrices $\widetilde{\mathbf{A}} \in \mathbb{C}^{(m-1)\times(m-1)}$ and $\widetilde{\mathbf{C}} \in \mathbb{C}^{l\times(m-1)}$ become one-dimension smaller. Therefore, we can apply the induction hypothesis to estimate $\widetilde{\mathbf{x}}$.

We can also write $\widetilde{\mathbf{A}}$ and $\widetilde{\mathbf{C}}$ in the same way that we write $\mathbf{A}$ and $\mathbf{C}$ as (2.35), (2.36) and (2.37), and define the corresponding parameters shown in (2.35), (2.36) and (2.37). To distinguish the parameters for $\widetilde{\mathbf{A}}$ and $\widetilde{\mathbf{C}}$ from the parameters for $\mathbf{A}$ and $\mathbf{C}$, we use tilde. For example, the dimension of $\mathbf{A}$ was $m \times m$, and we define the dimension of $\widetilde{\mathbf{A}}$ as $\widetilde{m} \times \widetilde{m}$. Likewise, the parameters $\widetilde{\mu}$, $\widetilde{\nu}_i$, $\widetilde{\lambda}_{i,j}$, $\widetilde{m}_{i,j}$, $\widetilde{p}_i$, $\widetilde{l}_i$ are defined for the system matrices $\widetilde{\mathbf{A}}$ and $\widetilde{\mathbf{C}}$ in the same ways as (2.35), (2.36) and (2.37).

By the induction hypothesis, we can find $\widetilde{m}'_1, \cdots, \widetilde{m}'_{\widetilde{\mu}} \in \mathbb{N}$, positive polynomials $\widetilde{p}_1(k), \cdots, \widetilde{p}_{\widetilde{\mu}}(k)$ and families of stopping times $\{\widetilde{S}_1(\epsilon,k) : k \in \mathbb{Z}^+, 0 < \epsilon < 1\}, \cdots, \{\widetilde{S}_{\widetilde{\mu}}(\epsilon,k) : k \in \mathbb{Z}^+, 0 < \epsilon < 1\}$ such that for all $0 < \epsilon < 1$ there exist $\max\{\bar{S}(\epsilon,k), k_{th}\} \leq \widetilde{k}_1 < \cdots < \widetilde{k}_{\widetilde{m}'_1} \leq \widetilde{S}_1(\epsilon,k) < \widetilde{k}_{\widetilde{m}'_1+1} < \cdots < \widetilde{k}_{\sum_{1\leq i\leq\widetilde{\mu}} \widetilde{m}'_i} \leq \widetilde{S}_{\widetilde{\mu}}(\epsilon,k)$ and a $\widetilde{m} \times (\sum_{1\leq i\leq\widetilde{\mu}} \widetilde{m}'_i) l$ matrix $\widetilde{\mathbf{M}}$ satisfying the following conditions:
(i") $\beta[\widetilde{k}_i] = 1$ for $1 \leq i \leq \sum_{1\leq i\leq\widetilde{\mu}} \widetilde{m}'_i$

(ii") $\widetilde{\mathbf{M}} \begin{bmatrix} \widetilde{\mathbf{C}}\widetilde{\mathbf{A}}^{-\widetilde{k}_1} \\ \widetilde{\mathbf{C}}\widetilde{\mathbf{A}}^{-\widetilde{k}_2} \\ \vdots \\ \widetilde{\mathbf{C}}\widetilde{\mathbf{A}}^{-\widetilde{k}_{\sum_{1 \le i \le \widetilde{\mu}} \widetilde{m}'_i}} \end{bmatrix} = \mathbf{I}$

(iii") $|\widetilde{\mathbf{M}}|_{max} \le \max_{1 \le i \le \widetilde{\mu}} \left\{ \frac{\widetilde{p}_i(\tilde{S}_i(\epsilon,k))}{\epsilon} |\widetilde{\lambda}_{i,1}|^{\widetilde{S}_i(\epsilon,k)} \right\}$

(iv") $\lim_{\epsilon \downarrow 0} \exp \limsup_{s \to \infty} \operatorname{ess\,sup} \frac{1}{s} \log \mathbb{P}\{\widetilde{S}_i(\epsilon,k) - \max\{\bar{S}(\epsilon,k), k_{th}\} = s | \mathcal{F}_{\bar{S}(\epsilon,k)}\} = \max_{1 \le j \le i} \left\{ p_e^{\frac{\widetilde{l}_j}{\widetilde{p}_j}} \right\}$ for $1 \le i \le \widetilde{\mu}$

(v") $\lim_{\epsilon \downarrow 0} \exp \limsup_{s \to \infty} \operatorname{ess\,sup} \frac{1}{s} \log \mathbb{P}\{\widetilde{S}_a(\epsilon,k) - \widetilde{S}_b(\epsilon,k) = s | \mathcal{F}_{\widetilde{S}_b(\epsilon,k)}\} \le \max_{b < i \le a} \left\{ p_e^{\frac{\widetilde{l}_i}{\widetilde{p}_i}} \right\}$ for $1 \le b < a \le \widetilde{\mu}$. Compared to Lemma 2.3, we can notice that the condition (iv") is slightly different from the condition (iv) of Lemma 2.3. The sup over k of (iv) in Lemma 2.3 is replaced by the ess sup. However, if we remind that $\max\{\bar{S}(\epsilon,k), k_{th}\}$ is a constant conditioned on[4] $\mathcal{F}_{\bar{S}(\epsilon,k)}$, we just replaced k of Lemma 2.3 with $\max\{\bar{S}(\epsilon,k), k_{th}\}$.

[4] More precise notations for $\widetilde{S}_1(\epsilon,k), \cdots, \widetilde{S}_\mu(\epsilon,k)$ are $\widetilde{S}_1(\epsilon, \max\{\bar{S}(\epsilon,k), k_{th}\}), \cdots, \widetilde{S}_\mu(\epsilon, \max\{\bar{S}(\epsilon,k), k_{th}\})$ since $\max\{\bar{S}(\epsilon,k), k_{th}\}$ plays the role of k of Lemma 2.3 after conditioning. However, we use the notation of the chapter for simplicity.

Here, we have

$$
\begin{aligned}
\widetilde{\mathbf{x}} &= \widetilde{\mathbf{M}} \begin{bmatrix} \widetilde{\mathbf{C}}\widetilde{\mathbf{A}}^{-\widetilde{k}_1} \\ \widetilde{\mathbf{C}}\widetilde{\mathbf{A}}^{-\widetilde{k}_2} \\ \vdots \\ \widetilde{\mathbf{C}}\widetilde{\mathbf{A}}^{-\widetilde{k}_{\sum_{1\le i\le\widetilde{\mu}} \widetilde{m}'_i}} \end{bmatrix} \widetilde{\mathbf{x}} \\
&= \widetilde{\mathbf{M}} \begin{bmatrix} \widetilde{\mathbf{C}}\widetilde{\mathbf{A}}^{-\widetilde{k}_1}\widetilde{\mathbf{x}} \\ \widetilde{\mathbf{C}}\widetilde{\mathbf{A}}^{-\widetilde{k}_2}\widetilde{\mathbf{x}} \\ \vdots \\ \widetilde{\mathbf{C}}\widetilde{\mathbf{A}}^{-\widetilde{k}_{\sum_{1\le i\le\widetilde{\mu}} \widetilde{m}'_i}}\widetilde{\mathbf{x}} \end{bmatrix} \\
&= \widetilde{\mathbf{M}} \begin{bmatrix} \mathbf{C}\mathbf{A}^{-\widetilde{k}_1}\mathbf{x} - \mathbf{R}(\widetilde{k}_1)(\mathbf{x})_{m_{1,1}} \\ \mathbf{C}\mathbf{A}^{-\widetilde{k}_2}\mathbf{x} - \mathbf{R}(\widetilde{k}_2)(\mathbf{x})_{m_{1,1}} \\ \vdots \\ \mathbf{C}\mathbf{A}^{-\widetilde{k}_{\sum_{1\le i\le\widetilde{\mu}} \widetilde{m}'_i}}\mathbf{x} - \mathbf{R}(\widetilde{k}_{\sum_{1\le i\le\widetilde{\mu}} \widetilde{m}'_i})(\mathbf{x})_{m_{1,1}} \end{bmatrix} (\because (7.98)) \\
&= \widetilde{\mathbf{M}} \left(\begin{bmatrix} \mathbf{C}\mathbf{A}^{-\widetilde{k}_1} \\ \mathbf{C}\mathbf{A}^{-\widetilde{k}_2} \\ \vdots \\ \mathbf{C}\mathbf{A}^{-\widetilde{k}_{\sum_{1\le i\le\widetilde{\mu}} \widetilde{m}'_i}} \end{bmatrix} \mathbf{x} - \begin{bmatrix} \mathbf{R}(\widetilde{k}_1) \\ \mathbf{R}(\widetilde{k}_2) \\ \vdots \\ \mathbf{R}(\widetilde{k}_{\sum_{1\le i\le\widetilde{\mu}} \widetilde{m}'_i}) \end{bmatrix} (\mathbf{x})_{m_{1,1}} \right) \\
&= \widetilde{\mathbf{M}} \left(\begin{bmatrix} \mathbf{C}\mathbf{A}^{-\widetilde{k}_1} \\ \mathbf{C}\mathbf{A}^{-\widetilde{k}_2} \\ \vdots \\ \mathbf{C}\mathbf{A}^{-\widetilde{k}_{\sum_{1\le i\le\widetilde{\mu}} \widetilde{m}'_i}} \end{bmatrix} \mathbf{x} - \begin{bmatrix} \mathbf{R}(\widetilde{k}_1) \\ \mathbf{R}(\widetilde{k}_2) \\ \vdots \\ \mathbf{R}(\widetilde{k}_{\sum_{1\le i\le\widetilde{\mu}} \widetilde{m}'_i}) \end{bmatrix} \bar{\mathbf{M}} \begin{bmatrix} \mathbf{C}\mathbf{A}^{-\bar{k}_1} \\ \mathbf{C}\mathbf{A}^{-\bar{k}_2} \\ \vdots \\ \mathbf{C}\mathbf{A}^{-\bar{k}_{\bar{m}}} \end{bmatrix} \mathbf{x} \right) (\because \text{the condition (ii) of Claim 7.8}) \\
&= \widetilde{\mathbf{M}} \begin{bmatrix} - \begin{bmatrix} \mathbf{R}(\widetilde{k}_1) \\ \mathbf{R}(\widetilde{k}_2) \\ \vdots \\ \mathbf{R}(\widetilde{k}_{\sum_{1\le i\le\widetilde{\mu}} \widetilde{m}'_i}) \end{bmatrix} \bar{\mathbf{M}} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{C}\mathbf{A}^{-\bar{k}_1} \\ \vdots \\ \mathbf{C}\mathbf{A}^{-\bar{k}_{\bar{m}}} \\ \mathbf{C}\mathbf{A}^{-\widetilde{k}_1} \\ \vdots \\ \mathbf{C}\mathbf{A}^{-\widetilde{k}_{\sum_{1\le i\le\widetilde{\mu}} \widetilde{m}'_i}} \end{bmatrix} \mathbf{x}. \qquad (7.99)
\end{aligned}
$$

When $|\lambda_{1,1}| > 1$, we have

$$\left| \widetilde{\mathbf{M}} \begin{bmatrix} - \begin{bmatrix} \mathbf{R}(\widetilde{k}_1) \\ \mathbf{R}(\widetilde{k}_2) \\ \vdots \\ \mathbf{R}(\widetilde{k}_{\sum_{1\le i\le\widetilde{\mu}} \widetilde{m}'_i}) \end{bmatrix} \bar{\mathbf{M}} & \mathbf{I} \end{bmatrix} \right|_{max}$$
$$\lesssim |\widetilde{\mathbf{M}}|_{max} \cdot \max \left\{ \left| \begin{bmatrix} \mathbf{R}(\widetilde{k}_1) \\ \mathbf{R}(\widetilde{k}_2) \\ \vdots \\ \mathbf{R}(\widetilde{k}_{\sum_{1\le i\le\widetilde{\mu}} \widetilde{m}'_i}) \end{bmatrix} \right|_{max} \left|\bar{\mathbf{M}}\right|_{max}, 1 \right\}$$
$$\lesssim \max_{1\le i\le\widetilde{\mu}} \left\{ \frac{\widetilde{p}_i(\widetilde{S}_i(\epsilon,k))}{\epsilon} |\widetilde{\lambda}_{i,1}|^{\widetilde{S}_i(\epsilon,k)} \right\} \cdot \max \left\{ \widetilde{p}(\widetilde{k}_1)|\lambda_{1,1}|^{-\widetilde{k}_1} \frac{\bar{p}\left(\bar{S}(\epsilon,k)\right)}{\epsilon} |\lambda_{1,1}|^{\bar{S}(\epsilon,k)}, 1 \right\} \tag{7.100}$$

where the last inequality follows from (iii"), $|\mathbf{R}(k)| \le \widetilde{p}(k)|\lambda_{1,1}|^{-k}$, $k_{th} \le \widetilde{k}_i$, and condition (iii) of Claim 7.8. Moreover, since $\bar{S}(\epsilon,k) \le \widetilde{k}_1 \le \widetilde{S}_i(\epsilon,k)$, there exist some positive polynomials $p'_i(k)$ such that

$$(7.100) \lesssim \max_{1\le i\le\widetilde{\mu}} \left\{ \frac{p'_i(\widetilde{S}_i(\epsilon,k))}{\epsilon^2} |\widetilde{\lambda}_{i,1}|^{\widetilde{S}_i(\epsilon,k)} \right\} \tag{7.101}$$

When $|\lambda_{1,1}| = 1$, $|\widetilde{\lambda}_{1,1}|$ is also 1. Thus, we have

$$\left| \widetilde{\mathbf{M}} \begin{bmatrix} - \begin{bmatrix} \mathbf{R}(\widetilde{k}_1) \\ \mathbf{R}(\widetilde{k}_2) \\ \vdots \\ \mathbf{R}(\widetilde{k}_{\sum_{1\le i\le\widetilde{\mu}} \widetilde{m}'_i}) \end{bmatrix} \bar{\mathbf{M}} & \mathbf{I} \end{bmatrix} \right|_{max}$$
$$\lesssim |\widetilde{\mathbf{M}}|_{max} \cdot \max \left\{ \left| \begin{bmatrix} \mathbf{R}(\widetilde{k}_1) \\ \mathbf{R}(\widetilde{k}_2) \\ \vdots \\ \mathbf{R}(\widetilde{k}_{\sum_{1\le i\le\widetilde{\mu}} \widetilde{m}'_i}) \end{bmatrix} \right|_{max} \left|\bar{\mathbf{M}}\right|_{max}, 1 \right\}$$
$$\lesssim \max_{1\le i\le\widetilde{\mu}} \left\{ \frac{\widetilde{p}_1(\widetilde{S}_1(\epsilon,k))}{\epsilon} \right\} \cdot \max \left\{ \widetilde{p}(\widetilde{k}_{\sum_{1\le i\le\widetilde{\mu}} \widetilde{m}'_i} \frac{\bar{p}\left(\bar{S}(\epsilon,k)\right)}{\epsilon}, 1 \right\}$$
$$\lesssim \frac{p'(\widetilde{S}_{\widetilde{\mu}}(\epsilon,k))}{\epsilon^2} \tag{7.102}$$

for some polynomial $p'_{\widetilde{\mu}}(k)$.

Since we can reconstruct $\mathbf{x}$ from $\widetilde{\mathbf{x}}$ and $(\mathbf{x})_{m_{1,1}}$, we can say there exists $\mathbf{M}$ such that

$$\mathbf{M}\begin{bmatrix}\mathbf{C}\mathbf{A}^{-\bar{k}_1}\\ \vdots \\ \mathbf{C}\mathbf{A}^{-\bar{k}_{\bar{m}}}\\ \mathbf{C}\mathbf{A}^{-\widetilde{k}_1}\\ \vdots \\ \mathbf{C}\mathbf{A}^{-\widetilde{k}_{\sum_{1\leq i\leq \widetilde{\mu}}\widetilde{m}_i}}\end{bmatrix}=\mathbf{I}.$$

By condition (ii) of Claim 7.8 and (7.99), such $\mathbf{M}$ satisfies the following:

$$|\mathbf{M}|_{max}\leq \max\left\{\left|\bar{\mathbf{M}}\right|_{max}, \left|\widetilde{\mathbf{M}}\begin{bmatrix}-\begin{bmatrix}\mathbf{R}(\widetilde{k}_1)\\ \mathbf{R}(\widetilde{k}_2)\\ \vdots \\ \mathbf{R}(\widetilde{k}_{\sum_{1\leq i\leq\widetilde{\mu}}\widetilde{m}'_i})\end{bmatrix}\bar{\mathbf{M}} & \mathbf{I}\end{bmatrix}\right|_{max}\right\}$$

$$\lesssim \max\left\{\frac{\bar{p}(\bar{S}(\epsilon,k))}{\epsilon}|\lambda_{1,1}|^{\bar{S}(\epsilon,k)}, \max_{1\leq i\leq\widetilde{\mu}}\left\{\frac{p'_i(\widetilde{S}_i(\epsilon,k))}{\epsilon^2}|\widetilde{\lambda}_{i,1}|^{\widetilde{S}_i(\epsilon,k)}\right\}\right\} \tag{7.103}$$

$$\leq \frac{1}{\epsilon^2}\max\left\{\bar{p}(\bar{S}(\epsilon,k))|\lambda_{1,1}|^{\bar{S}(\epsilon,k)}, \max_{1\leq i\leq\widetilde{\mu}}\left\{p'_i(\widetilde{S}_i(\epsilon,k))|\widetilde{\lambda}_{i,1}|^{\widetilde{S}_i(\epsilon,k)}\right\}\right\}. \tag{7.104}$$

Here, (7.103) follows from the condition (iii) of Claim 7.8, (7.101), (7.102).

Moreover, since k_{th} is a constant, the condition (iv) of Claim 7.8 implies

$$\lim_{\epsilon\downarrow 0}\exp\limsup_{s\to\infty}\sup_{k\in\mathbb{Z}^+}\frac{1}{s}\log\mathbb{P}\left\{\max\left\{\bar{S}(\epsilon,k),k_{th}\right\}-k=s\right\}=p_e^{\frac{l_1}{p_1}}. \tag{7.105}$$

Therefore, by applying Lemma 7.2 together with (7.105) and (iv") we get

$$\lim_{\epsilon\downarrow 0}\exp\limsup_{s\to\infty}\sup_{k\in\mathbb{Z}^+}\frac{1}{s}\log\mathbb{P}\{\widetilde{S}_i(\epsilon,k)-k=s\}=\max\left\{p_e^{\frac{l_1}{p_1}},\max_{1\leq j\leq i}\left\{p_e^{\frac{\widetilde{l}_j}{\widetilde{p}_j}}\right\}\right\}. \tag{7.106}$$

We finish the proof by dividing into two cases depending on $\widetilde{\mu}$. Since $\widetilde{\mathbf{A}}$ is obtained by erasing just one row and column of $\mathbf{A}$, the relation between $\widetilde{\mu}$ and μ is either $\widetilde{\mu}=\mu$ or $\widetilde{\mu}=\mu-1$.

(1) When $\widetilde{\mu}=\mu$.

In this case, the number of the eigenvalue cycles remains the same. We can see that $|\widetilde{\lambda}_{i,1}|=|\lambda_{i,1}|$. $\mathbf{A_1}$ and $\widetilde{\mathbf{A}}_\mathbf{1}$ may be the same or $\widetilde{\mathbf{A}}_\mathbf{1}$ has smaller dimension than $\mathbf{A_1}$. Thus, the new system $\widetilde{\mathbf{A}}_\mathbf{1}$ becomes easier to estimate, and $\frac{\widetilde{l}_1}{\widetilde{p}_1}\geq\frac{l_1}{p_1}$, i.e. $p_e^{\frac{\widetilde{l}_1}{\widetilde{p}_1}}\leq p_e^{\frac{l_1}{p_1}}$. $\mathbf{A_i}$ and $\widetilde{\mathbf{A}}_\mathbf{i}$ are the same for all $2\leq i\leq\mu$, so $\frac{\widetilde{l}_i}{\widetilde{p}_i}=\frac{l_i}{p_i}$ for $2\leq j\leq\mu$. Define $S_i(\epsilon^2,k):=\widetilde{S}_i(\epsilon,k)$, $p_1(k):=\bar{p}(k)+p'_1(k)$, and $p_i(k):=p'_i(k)$ for $2\leq i\leq\mu$. Then, (7.104), (7.106) and (v") reduces as follows:

$$|\mathbf{M}|_{max}\leq\max_{1\leq i\leq\mu}\left\{\frac{p_i(S_i(\epsilon,k))}{\epsilon}|\lambda_{i,1}|^{S_i(\epsilon,k)}\right\},$$

$$\lim_{\epsilon\downarrow 0}\exp\limsup_{s\to\infty}\sup_{k\in\mathbb{Z}^+}\frac{1}{s}\log\mathbb{P}\{S_i(\epsilon,k)-k=s\}\le\max_{1\le j\le i}\left\{p_e^{\frac{l_j}{p_j}}\right\},$$

$$\lim_{\epsilon\downarrow 0}\exp\limsup_{s\to\infty}\operatorname{ess\,sup}\frac{1}{s}\log\mathbb{P}\{S_a(\epsilon,k)-S_b(\epsilon,k)=s|\mathcal{F}_{S_b(\epsilon,k)}\}\le\max_{b<i\le a}\left\{p_e^{\frac{l_i}{p_i}}\right\}.$$

Here, we reparametrized ϵ^2 to ϵ. Therefore, the lemma is true for this case.

(2) When $\widetilde{\mu}=\mu-1$.

Since one eigenvalue cycle has disappeared, we can see that $|\widetilde{\lambda}_{1,1}|=|\lambda_{2,1}|,|\widetilde{\lambda}_{2,1}|=|\lambda_{3,1}|,\cdots,|\widetilde{\lambda}_{\widetilde{\mu},1}|=|\lambda_{\mu,1}|$. Moreover, $\widetilde{\mathbf{A}}_{\mathbf{i}}=\mathbf{A}_{\mathbf{i+1}}$ for $1\le i\le\widetilde{\mu}$ and $\frac{\widetilde{l}_i}{\widetilde{p}_i}=\frac{l_{i+1}}{p_{i+1}}$ for $1\le i\le\widetilde{\mu}$. Define $S_1(\epsilon^2,k):=\bar{S}(\epsilon,k)$, $p_1(k):=\bar{p}(k)$, $S_i(\epsilon^2,k):=\widetilde{S}_{i-1}(\epsilon,k)$ and $p_i(k):=p'_{i-1}(k)$ for $2\le i\le\mu$. We will also reparametrize ϵ^2 to ϵ. Then, (7.104) reduces to

$$|\mathbf{M}|_{max}\le\max_{1\le i\le\mu}\left\{\frac{p_i(S_i(\epsilon,k))}{\epsilon}|\lambda_{i,1}|^{S_i(\epsilon,k)}\right\}.$$

By the definition of $S_1(\epsilon,k)$, the condition (iv) of Claim 7.8 reduces to

$$\lim_{\epsilon\downarrow 0}\exp\limsup_{s\to\infty}\sup_{k\in\mathbb{Z}^+}\frac{1}{s}\log\mathbb{P}\{S_1(\epsilon,k)-k=s\}\le p_e^{\frac{l_1}{p_1}}.$$

By (7.106) and the definition of $S_i(\epsilon,k)$, we have for all $2\le i\le\mu$,

$$\lim_{\epsilon\downarrow 0}\exp\limsup_{s\to\infty}\sup_{k\in\mathbb{Z}^+}\frac{1}{s}\log\mathbb{P}\{S_i(\epsilon,k)-k=s\}\le\max\left\{p_e^{\frac{l_1}{p_1}},\max_{1\le j\le i-1}\left\{p_e^{\frac{\widetilde{l}_j}{\widetilde{p}_j}}\right\}\right\}=\max_{1\le j\le i}\left\{p_e^{\frac{l_i}{p_i}}\right\}.$$

By (iv"), (v") and the definition of $S_i(\epsilon,k)$, we have for all $1\le b<a\le\mu$,

$$\lim_{\epsilon\downarrow 0}\exp\limsup_{s\to\infty}\operatorname{ess\,sup}\frac{1}{s}\log\mathbb{P}\{S_a(\epsilon,k)-S_b(\epsilon,k)=s|\mathcal{F}_{S_b(\epsilon,k)}\}\le\max_{b<i\le a}\left\{p_e^{\frac{l_i}{p_i}}\right\}.$$

Therefore, the lemma is also true for this case.

Thus, the proof is finished. □

Chapter 8

Appendix for Chapter 3

8.1 Network Linearization for General Information Flow

In this section, we will extend the network linearization idea of Section 3.2.2 from the point-to-point case to general information flow cases – multicast, broadcast and multiple-unicast. The main idea for this generalization is the relationship between network linearization and control over LTI networks discussed in Section 3.6.

8.1.1 Multicast

From the discussion in the point-to-point case, we can expect that to linearize multicast problems, we have to introduce circulation arcs in a way that corresponds with Fig. 3.4. Fig. 8.1 shows how the circulation arc has to be introduced. One circulation arc (which corresponds to an unstable plant as discussed in Section 3.6.1) is connected to both receivers.

We will essentially use the same notation and assumptions as Section 3.2.2. Let the one-transmitter two-receiver LTI network of Fig. 8.1 without circulation arcs be $\mathcal{N}_{mul}(z)$. Denote the dimension of Y_1 as d_{rx1} and Y_2 as d_{rx2}. Let the transfer function from the transmitter to the receiver 1 of $\mathcal{N}_{mul}(z)$ be $G_{tx,rx1}(z,K)$, and the transfer function from the transmitter to the receiver 2 be $G_{tx,rx2}(z,K)$. Here, the transfer function can be computed in the same way as Theorem 3.1.

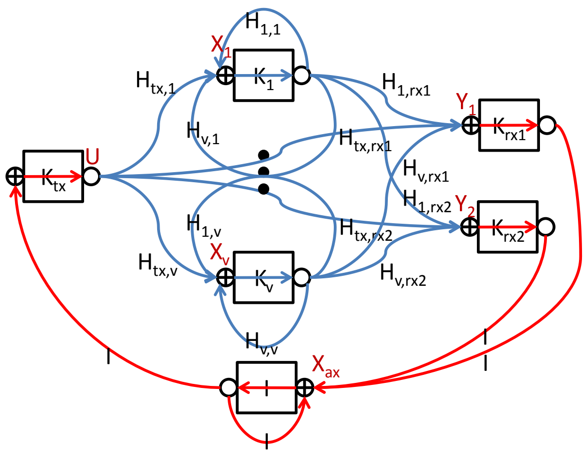


Figure 8.1: Multicast LTI network $\mathcal{N}_{mul}(Z)$ with circulation arc added in

Then, similar to Section 3.2.2, the following relation has to hold:

$$\begin{bmatrix} X_{ax} \\ Y_1 \\ Y_2 \\ X_1 \\ \vdots \\ X_v \end{bmatrix} = \begin{bmatrix} I & K_{rx1} & K_{rx2} & 0 & \cdots & 0 \\ H_{tx,rx1}K_{tx} & 0 & 0 & H_{1,rx1}K_1 & \cdots & H_{v,rx1}K_v \\ H_{tx,rx2}K_{tx} & 0 & 0 & H_{1,rx2}K_1 & \cdots & H_{v,rx2}K_v \\ H_{tx,1}K_{tx} & 0 & 0 & H_{1,1}K_1 & \cdots & H_{v,1}K_v \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ H_{tx,v}K_{tx} & 0 & 0 & H_{1,v}K_1 & \cdots & H_{v,v}K_v \end{bmatrix} \begin{bmatrix} X_{ax} \\ Y_1 \\ Y_2 \\ X_1 \\ \vdots \\ X_v \end{bmatrix}$$

$$(\Leftrightarrow) \underbrace{\begin{bmatrix} 0 & -K_{rx1} & -K_{rx2} & 0 & \cdots & 0 \\ -H_{tx,rx1}K_{tx} & I & 0 & -H_{1,rx1}K_1 & \cdots & -H_{v,rx1}K_v \\ -H_{tx,rx2}K_{tx} & 0 & I & -H_{1,rx2}K_1 & \cdots & -H_{v,rx2}K_v \\ -H_{tx,1}K_{tx} & 0 & 0 & I-H_{1,1}K_1 & \cdots & -H_{v,1}K_v \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ -H_{tx,v}K_{tx} & 0 & 0 & -H_{1,v}K_1 & \cdots & I-H_{v,v}K_v \end{bmatrix}}_{:=G_{lin}(z,K)} \begin{bmatrix} X_{ax} \\ Y_1 \\ Y_2 \\ X_1 \\ \vdots \\ X_v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Then, we have

$$G_{lin}(z,K) = \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & I & 0 & 0 & \cdots & 0 \\ 0 & 0 & I & 0 & \cdots & 0 \\ 0 & 0 & 0 & I & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & 0 & 0 & \cdots & I \end{bmatrix}}_{:=A} + \underbrace{\begin{bmatrix} 0 \\ H_{tx,rx1} \\ H_{tx,rx2} \\ H_{tx,1} \\ \vdots \\ H_{tx,v} \end{bmatrix}}_{:=B_{tx}} K_{tx} \underbrace{\begin{bmatrix} -I & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}}_{:=C_{tx}}$$

$$+ \underbrace{\begin{bmatrix} 0 \\ H_{1,rx1} \\ H_{1,rx2} \\ H_{1,1} \\ \vdots \\ H_{1,v} \end{bmatrix}}_{:=B_1} K_1 \underbrace{\begin{bmatrix} 0 & 0 & 0 & -I & \cdots & 0 \end{bmatrix}}_{:=C_1} + \cdots + \underbrace{\begin{bmatrix} 0 \\ H_{v,rx1} \\ H_{v,rx2} \\ H_{v,1} \\ \vdots \\ H_{v,v} \end{bmatrix}}_{:=B_v} K_v \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 & \cdots & -I \end{bmatrix}}_{:=C_v}$$

$$+ \underbrace{\begin{bmatrix} I \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{:=B_{rx1}} K_{rx1} \underbrace{\begin{bmatrix} 0 & -I & 0 & 0 & \cdots & 0 \end{bmatrix}}_{:=C_{rx1}} + \underbrace{\begin{bmatrix} I \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{:=B_{rx2}} K_{rx2} \underbrace{\begin{bmatrix} 0 & 0 & -I & 0 & \cdots & 0 \end{bmatrix}}_{:=C_{rx2}}$$

Let

$$G_{tx',rx1'}(z,K) := A + B_{tx}K_{tx}C_{tx} + B_{rx1}K_{rx1}C_{rx1} + \sum_{1 \le i \le v} B_i K_i C_i$$

$$G_{tx',rx2'}(z,K) := A + B_{tx}K_{tx}C_{tx} + B_{rx2}K_{rx2}C_{rx2} + \sum_{1 \le i \le v} B_i K_i C_i$$

and $d := \dim \begin{bmatrix} Y_1 \\ Y_2 \\ X_1 \\ \vdots \\ X_v \end{bmatrix}$. Let $\mathcal{N}_{lin}^{mul}(z)$ be the network shown in Fig. 8.2. Then, we can easily see $G_{tx',rx1'}(z,K)$ is the transfer function from tx' to rx'_1 of $\mathcal{N}_{mul}^{lin}(z)$, and $G_{tx',rx2'}(z,K)$ is the transfer function from tx' to rx'_2 of $\mathcal{N}_{mul}^{lin}(z)$.

Then, like Section 3.2.2 we can show the equivalence between $\mathcal{N}_{mul}(z)$ and $\mathcal{N}_{mul}^{lin}(z)$.

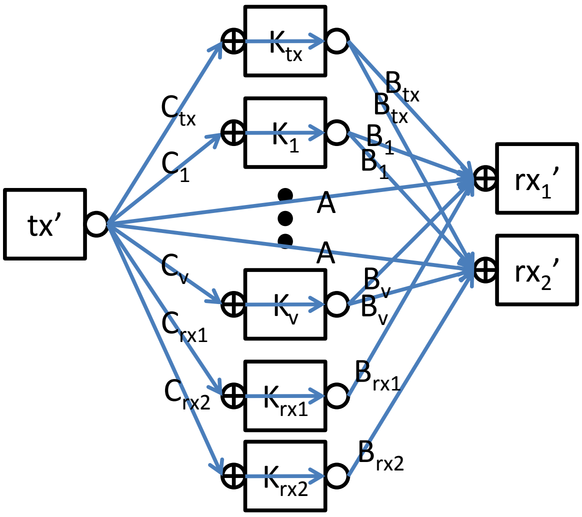


Figure 8.2: Linearized LTI network of Multicast problem, $\mathcal{N}_{mul}^{lin}(z)$

Theorem 8.1. *Let* $K_{tx} \in \mathbb{F}[z]^{d_{tx} \times d_{ax}}$*,* $K_i \in \mathbb{F}[z]^{d_{i,in} \times d_{i,out}}$*,* $K_{rx1} \in \mathbb{F}[z]^{d_{ax} \times d_{rx1}}$ *and* $K_{rx2} \in \mathbb{F}[z]^{d_{ax} \times d_{rx2}}$*. We also assume that*

$$\begin{bmatrix} I - H_{1,1}K_1 & \cdots & -H_{v,1}K_v \\ \vdots & \ddots & \vdots \\ -H_{1,v}K_1 & \cdots & I - H_{v,v}K_v \end{bmatrix} \text{ is invertible.}$$

Then, for all $d_1, d_2 \in \mathbb{Z}^+$

$$(i)\,\text{rank}(K_{rx1}(z)G_{tx,rx1}(z, K(z))K_{tx}(z)) \geq d_1$$
$$(ii)\,\text{rank}(K_{rx2}(z)G_{tx,rx2}(z, K(z))K_{tx}(z)) \geq d_2$$

if and only if

$$(a)\,\text{rank}\, G_{tx',rx1'}(z, K(z)) \geq d + d_1$$
$$(b)\,\text{rank}\, G_{tx',rx2'}(z, K(z)) \geq d + d_2$$

Proof. Similar to Lemma 3.3. □

Remark 1. The result of this theorem can be easily generalized to multiple receivers, which we omit for simplicity.

Remark 2. To apply this theorem to multicast problems and send a message with rate r, we can simply put $d_1 = d_2 = r$. Moreover, just as we did in Figure 3.7, the condition that $\begin{bmatrix} I - H_{1,1}K_1 & \cdots & -H_{v,1}K_v \\ \vdots & \ddots & \vdots \\ -H_{1,v}K_1 & \cdots & I - H_{v,v}K_v \end{bmatrix}$ is invertible can be included as a part of the communication

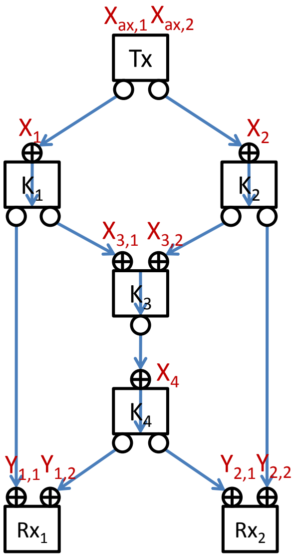


Figure 8.3: Butterfly Example for Multicast. The gains of all edges are 1.

problem by introducing an additional receiver. Following the similar procedure of Section 3.2.3, we can design an LTI multicast scheme.

Remark 3. Fig. 8.3 and Fig. 8.4 shows the famous butterfly example in network coding [1] and its corresponding linearized network. Here, we can see the linearized network has more input and output vertices, but is **topologically** simpler — a single-hop multicast network. Because there are no cycles, the additional receiver in Remark 2 is not required.

8.1.2 Broadcast

Inspired by Figure 3.20, we introduce circulation arcs as shown in Figure 8.5 to linearize broadcast problems. We introduce two circulation arcs which correspond to the two unstable plants of Figure 3.20, and the two circulation arcs are connected to different receivers as two plants are controlled by different controllers in Figure 3.20.

We basically use the same notations and assumptions of the previous section. Let the one-transmitter two-receiver LTI network of Fig. 8.5 without circulation arcs be $\mathcal{N}_{br}(z)$. Denote the dimension of X_{ax1} as d_{ax1} and X_{ax2} as d_{ax2}. Then, as we can see from the figure, K_{tx1} is a $d_{tx} \times d_{ax1}$ matrix and K_{tx2} is a $d_{tx} \times d_{ax2}$ matrix. Let the transfer function from the transmitter to the receiver 1 of $\mathcal{N}_{mul}(z)$ be $G_{tx,rx1}(z, K)$, and the transfer function from the transmitter to the receiver 2 be $G_{tx,rx2}(z, K)$.

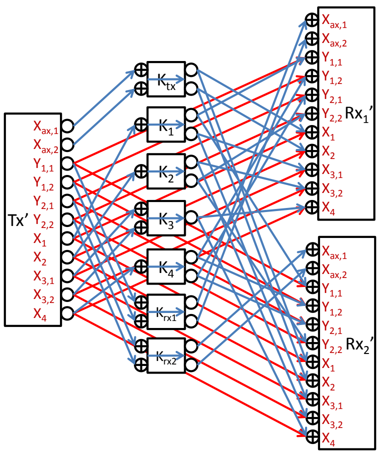


Figure 8.4: Linearized Network for Butterfly Example of Fig. 8.3. The gain of each edge from Tx' to K_{tx}, K_i, K_{rx1}, K_{rx2} is -1, and the gains for the other edges are all 1.

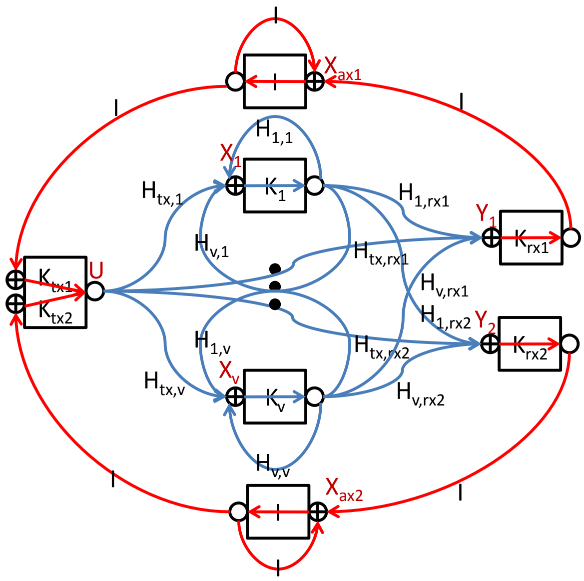


Figure 8.5: Broadcast LTI network $\mathcal{N}_{br}(z)$ with circulation arcs added in

Then, the following relation has to hold:

$$\begin{bmatrix} X_{ax1} \\ X_{ax2} \\ Y_1 \\ Y_2 \\ X_1 \\ \vdots \\ X_v \end{bmatrix} = \begin{bmatrix} I & 0 & K_{rx1} & 0 & 0 & \cdots & 0 \\ 0 & I & 0 & K_{rx2} & 0 & \cdots & 0 \\ H_{tx,rx1}K_{tx1} & H_{tx,rx1}K_{tx2} & 0 & 0 & H_{1,rx1}K_1 & \cdots & H_{v,rx1}K_v \\ H_{tx,rx2}K_{tx1} & H_{tx,rx2}K_{tx2} & 0 & 0 & H_{1,rx2}K_1 & \cdots & H_{v,rx2}K_v \\ H_{tx,1}K_{tx1} & H_{tx,1}K_{tx2} & 0 & 0 & H_{1,1}K_1 & \cdots & H_{v,1}K_v \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ H_{tx,v}K_{tx1} & H_{tx,v}K_{tx2} & 0 & 0 & H_{1,v}K_1 & \cdots & H_{v,v}K_v \end{bmatrix} \begin{bmatrix} X_{ax1} \\ X_{ax2} \\ Y_1 \\ Y_2 \\ X_1 \\ \vdots \\ X_v \end{bmatrix}$$

$$(\Leftrightarrow) \underbrace{\begin{bmatrix} 0 & 0 & -K_{rx1} & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & -K_{rx2} & 0 & \cdots & 0 \\ -H_{tx,rx1}K_{tx1} & -H_{tx,rx1}K_{tx2} & I & 0 & -H_{1,rx1}K_1 & \cdots & -H_{v,rx1}K_v \\ -H_{tx,rx2}K_{tx1} & -H_{tx,rx2}K_{tx2} & 0 & I & -H_{1,rx2}K_1 & \cdots & -H_{v,rx2}K_v \\ -H_{tx,1}K_{tx1} & -H_{tx,1}K_{tx2} & 0 & 0 & I - H_{1,1}K_1 & \cdots & -H_{v,1}K_v \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ -H_{tx,v}K_{tx1} & -H_{tx,v}K_{tx2} & 0 & 0 & -H_{1,v}K_1 & \cdots & I - H_{v,v}K_v \end{bmatrix}}_{:=G_{br}^{lin}(z,K)} \begin{bmatrix} X_{ax1} \\ X_{ax2} \\ Y_1 \\ Y_2 \\ X_1 \\ \vdots \\ X_v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Thus, we have

$$G_{br}^{lin}(z,K) = \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & I & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & I & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & I & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & I \end{bmatrix}}_{:=A} + \underbrace{\begin{bmatrix} 0 \\ 0 \\ H_{tx,rx1} \\ H_{tx,rx2} \\ H_{tx,1} \\ \vdots \\ H_{tx,v} \end{bmatrix}}_{:=B_{tx1}} K_{tx1} \underbrace{\begin{bmatrix} -I & 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}}_{:=C_{tx1}}$$

$$+ \underbrace{\begin{bmatrix} 0 \\ 0 \\ H_{tx,rx1} \\ H_{tx,rx2} \\ H_{tx,1} \\ \vdots \\ H_{tx,v} \end{bmatrix}}_{:=B_{tx2}} K_{tx2} \underbrace{\begin{bmatrix} 0 & -I & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}}_{:=C_{tx2}}$$

$$+ \underbrace{\begin{bmatrix} 0 \\ 0 \\ H_{1,rx1} \\ H_{1,rx2} \\ H_{1,1} \\ \vdots \\ H_{1,v} \end{bmatrix}}_{:=B_{1}} K_{1} \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 & -I & \cdots & 0 \end{bmatrix}}_{:=C_{1}} + \cdots + \underbrace{\begin{bmatrix} 0 \\ 0 \\ H_{v,rx1} \\ H_{v,rx2} \\ H_{v,1} \\ \vdots \\ H_{v,v} \end{bmatrix}}_{:=B_{v}} K_{v} \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \cdots & -I \end{bmatrix}}_{:=C_{v}}$$

$$+ \underbrace{\begin{bmatrix} I \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{:=B_{rx1}} K_{rx1} \underbrace{\begin{bmatrix} 0 & 0 & -I & 0 & 0 & \cdots & 0 \end{bmatrix}}_{:=C_{rx1}} + \underbrace{\begin{bmatrix} 0 \\ I \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{:=B_{rx2}} K_{rx2} \underbrace{\begin{bmatrix} 0 & 0 & 0 & -I & 0 & \cdots & 0 \end{bmatrix}}_{:=C_{rx2}}$$

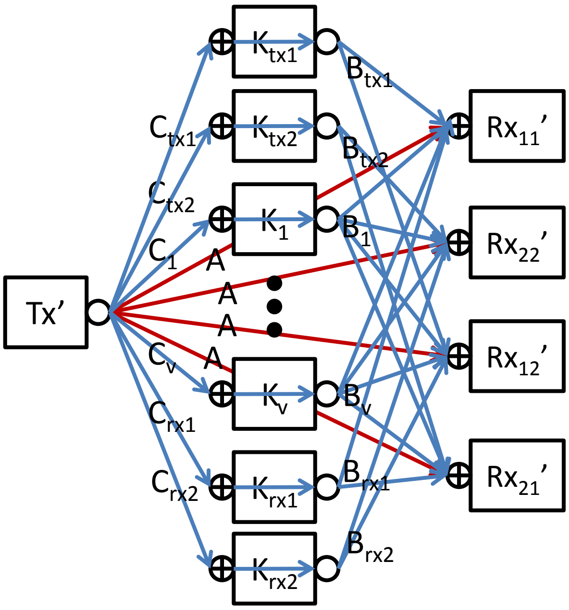


Figure 8.6: Linearized LTI network of a Broadcast problem, $\mathcal{N}_{br}^{lin}(z)$

Let

$$G_{tx',rx11'}(z,K) := A + B_{tx1}K_{tx1}C_{tx1} + B_{rx1}K_{rx1}C_{rx1} + \sum_{1\le i\le v} B_iK_iC_i \tag{8.1}$$

$$G_{tx',rx22'}(z,K) := A + B_{tx2}K_{tx2}C_{tx2} + B_{rx2}K_{rx2}C_{rx2} + \sum_{1\le i\le v} B_iK_iC_i$$

$$G_{tx',rx12'}(z,K) := A + B_{tx2}K_{tx2}C_{tx2} + B_{rx1}K_{rx1}C_{rx1} + \sum_{1\le i\le v} B_iK_iC_i$$

$$G_{tx',rx12'}(z,K) := A + B_{tx1}K_{tx1}C_{tx1} + B_{rx2}K_{rx2}C_{rx2} + \sum_{1\le i\le v} B_iK_iC_i$$

Let $\mathcal{N}_{lin}^{br}(z)$ be the network shown in Fig. 8.6. Then, we can easily see $G_{tx',rx11'}(z,K), \cdots, G_{tx',rx12'}(z,K)$ correspond to the transfer functions from tx' to $rx'_{11}, \cdots, rx'_{12}$ of $\mathcal{N}_{lin}^{br}(z)$ respectively.

Then, the relationship between $\mathcal{N}_{br}(z)$ and $\mathcal{N}_{br}^{lin}(z)$ is given as follows.

Theorem 8.2. *Let $K_{tx1}(z) \in \mathbb{F}[z]^{d_{tx}\times d_{ax1}}$, $K_{tx2}(z) \in \mathbb{F}[z]^{d_{tx}\times d_{ax2}}$, $K_i(z) \in \mathbb{F}[z]^{d_{i,in}\times d_{i,out}}$, $K_{rx1}(z) \in \mathbb{F}[z]^{d_{ax1}\times d_{rx1}}$ and $K_{rx2}(z) \in \mathbb{F}[z]^{d_{ax2}\times d_{rx2}}$. We also assume that*

$$\begin{bmatrix} I - H_{1,1}K_1 & \cdots & -H_{v,1}K_v \\ \vdots & \ddots & \vdots \\ -H_{1,v}K_1 & \cdots & I - H_{v,v}K_v \end{bmatrix} \text{ is invertible.}$$

Then, for all $d_1, d_2, d_3, d_4 \in \mathbb{Z}^+$*, the following two conditions are equivalent.*

$$(i)\,\text{rank}\,K_{rx1}(z)G_{tx,rx1}(z,K(z))K_{tx1}(z) \geq d_1$$
$$(ii)\,\text{rank}\,K_{rx2}(z)G_{tx,rx2}(z,K(z))K_{tx2}(z) \geq d_2$$
$$(iii)\,\text{rank}\,K_{rx2}(z)G_{tx,rx2}(z,K(z))K_{tx1}(z) \leq d_3$$
$$(iv)\,\text{rank}\,K_{rx1}(z)G_{tx,rx1}(z,K(z))K_{tx2}(z) \leq d_4$$

if and only if

$$(a)\,\text{rank}\,G_{tx',rx11'}(z,K(z)) \geq d + d_1$$
$$(b)\,\text{rank}\,G_{tx',rx22'}(z,K(z)) \geq d + d_2$$
$$(c)\,\text{rank}\,G_{tx',rx12'}(z,K(z)) \leq d + d_3$$
$$(d)\,\text{rank}\,G_{tx',rx21'}(z,K(z)) \leq d + d_4$$

Proof. Similar to Lemma 3.3. □

Remark 1. The result of this theorem can be easily generalized to multiple receivers. In three receiver case, we will see 9 conditions. For a general n receiver case, we will see n^2 conditions since each receiver will see n different signals (one desired signal and $n-1$ interference).

Remark 2. To design a broadcast scheme which communicates a message with rate r_1 to receiver 1 and at the same time another message with rate r_2 to receiver 2, we can choose the problem parameters as $d_1 = r_1, d_2 = r_2, d_3 = 0, d_4 = 0$. Any scheme which satisfies the condition $(a)-(d)$, and the existence condition of transfer functions can be immediately applied to the original problem and give a broadcast communication scheme.

Remark 3. The linearized network of Figure 8.6 can be understood as a two-receiver and two-eavesdropper secrecy problem. The receivers $rx11'$ and $rx22'$ want to receive $d+d_1$ and $d+d_2$ dimensional information about the messages (possibly, common) respectively. While at the same time, we do not want to give more than $d+d_3$ and $d+d_4$ dimensions about the message to the eavesdroppers $rx12'$ and $rx21'$.

The receivers $rx11'$ and $rx22'$ in the linearized network reflect that the desired messages have to be received in the original problem. The eavesdropper $rx12'$ and $rx21'$ in the linearized network reflects that the undesired messages must be removable in the original problem.

8.1.3 Multiple-Unicast

As the only difference between Figure 3.20 and Figure 3.21 is the observers, we introduce circulation arcs in the same way as the broadcast problems in Figure 8.5. Fig. 8.7 shows the multiple-unicast LTI network $\mathcal{N}_{uni}(z)$ with the circulation arcs.

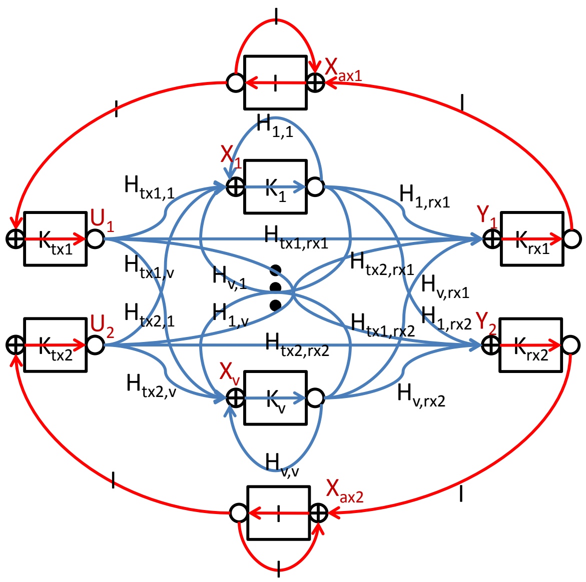


Figure 8.7: Multiple Unicast LTI network $\mathcal{N}_{uni}(z)$ with circulation arc added in

We essentially repeat the previous argument. Let's use the same notations and assumptions of the previous section. Denote the dimension of U_1, U_2, Y_1, Y_2 as $d_{tx1}, d_{tx2}, d_{rx1}, d_{rx2}$ respectively. The transfer functions between the transmitters and the receivers are denoted as $G_{tx1,rx1}(z,K)$, $G_{tx1,rx2}(z,K)$, $G_{tx2,rx1}(z,K)$, $G_{tx2,rx2}(z,K)$.

Then, we have the following relationship.

$$\begin{bmatrix} X_{ax1} \\ X_{ax2} \\ Y_1 \\ Y_2 \\ X_1 \\ \vdots \\ X_v \end{bmatrix} = \begin{bmatrix} I & 0 & K_{rx1} & 0 & 0 & \cdots & 0 \\ 0 & I & 0 & K_{rx2} & 0 & \cdots & 0 \\ H_{tx1,rx1}K_{tx1} & H_{tx2,rx1}K_{tx2} & 0 & 0 & H_{1,rx1}K_1 & \cdots & H_{v,rx1}K_v \\ H_{tx1,rx2}K_{tx1} & H_{tx2,rx2}K_{tx2} & 0 & 0 & H_{1,rx2}K_1 & \cdots & H_{v,rx2}K_v \\ H_{tx1,1}K_{tx1} & H_{tx2,1}K_{tx2} & 0 & 0 & H_{1,1}K_1 & \cdots & H_{v,1}K_v \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ H_{tx1,v}K_{tx1} & H_{tx2,v}K_{tx2} & 0 & 0 & H_{1,v}K_1 & \cdots & H_{v,v}K_v \end{bmatrix} \begin{bmatrix} X_{ax1} \\ X_{ax2} \\ Y_1 \\ Y_2 \\ X_1 \\ \vdots \\ X_v \end{bmatrix}$$

$$(\Leftrightarrow) \underbrace{\begin{bmatrix} 0 & 0 & -K_{rx1} & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & -K_{rx2} & 0 & \cdots & 0 \\ -H_{tx1,rx1}K_{tx1} & -H_{tx2,rx1}K_{tx2} & I & 0 & -H_{1,rx1}K_1 & \cdots & -H_{v,rx1}K_v \\ -H_{tx1,rx2}K_{tx1} & -H_{tx2,rx2}K_{tx2} & 0 & I & -H_{1,rx2}K_1 & \cdots & -H_{v,rx2}K_v \\ -H_{tx1,1}K_{tx1} & -H_{tx2,1}K_{tx2} & 0 & 0 & I-H_{1,1}K_1 & \cdots & -H_{v,1}K_v \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ -H_{tx1,v}K_{tx1} & -H_{tx2,v}K_{tx2} & 0 & 0 & -H_{1,v}K_1 & \cdots & I-H_{v,v}K_v \end{bmatrix}}_{:=G^{lin}_{uni}(z,K)} \begin{bmatrix} X_{ax1} \\ X_{ax2} \\ Y_1 \\ Y_2 \\ X_1 \\ \vdots \\ X_v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Therefore, we have

$$G^{lin}_{uni}(z,K) = \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & I & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & I & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & I & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & I \end{bmatrix}}_{:=A} + \underbrace{\begin{bmatrix} 0 \\ 0 \\ H_{tx1,rx1} \\ H_{tx1,rx2} \\ H_{tx1,1} \\ \vdots \\ H_{tx1,v} \end{bmatrix}}_{:=B_{tx1}} K_{tx1} \underbrace{\begin{bmatrix} -I & 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}}_{:=C_{tx1}}$$

$$+ \underbrace{\begin{bmatrix} 0 \\ 0 \\ H_{tx2,rx1} \\ H_{tx2,rx2} \\ H_{tx2,1} \\ \vdots \\ H_{tx2,v} \end{bmatrix}}_{:=B_{tx2}} K_{tx2} \underbrace{\begin{bmatrix} 0 & -I & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}}_{:=C_{tx2}}$$

$$+ \underbrace{\begin{bmatrix} 0 \\ 0 \\ H_{1,rx1} \\ H_{1,rx2} \\ H_{1,1} \\ \vdots \\ H_{1,v} \end{bmatrix}}_{:=B_{1}} K_{1} \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 & -I & \cdots & 0 \end{bmatrix}}_{:=C_{1}} + \cdots + \underbrace{\begin{bmatrix} 0 \\ 0 \\ H_{v,rx1} \\ H_{v,rx2} \\ H_{v,1} \\ \vdots \\ H_{v,v} \end{bmatrix}}_{:=B_{v}} K_{v} \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \cdots & -I \end{bmatrix}}_{:=C_{v}}$$

$$+ \underbrace{\begin{bmatrix} I \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{:=B_{rx1}} K_{rx1} \underbrace{\begin{bmatrix} 0 & 0 & -I & 0 & 0 & \cdots & 0 \end{bmatrix}}_{:=C_{rx1}} + \underbrace{\begin{bmatrix} 0 \\ I \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{:=B_{rx2}} K_{rx2} \underbrace{\begin{bmatrix} 0 & 0 & 0 & -I & 0 & \cdots & 0 \end{bmatrix}}_{:=C_{rx2}}$$

Use the same definitions of (8.1) for $G_{tx1,rx1}(z,K), \cdots, G_{tx2,rx2}(z,K)$. These transfer functions are the transfer functions of $\mathcal{N}^{lin}_{uni}(z)$ as before.

Then, Theorem 8.2 essentially holds for multiple unicast problems as well.

Theorem 8.3. *Let* $K_{tx1}(z) \in \mathbb{F}[z]^{d_{tx1}\times d_{ax1}}$, $K_{tx2}(z) \in \mathbb{F}[z]^{d_{tx2}\times d_{ax2}}$, $K_i(z) \in \mathbb{F}[z]^{d_{i,in}\times d_{i,out}}$, $K_{rx1}(z) \in \mathbb{F}[z]^{d_{ax1}\times d_{rx1}}$ *and* $K_{rx2}(z) \in \mathbb{F}[z]^{d_{ax2}\times d_{rx2}}$. *We also assume that*

$$\begin{bmatrix} I - H_{1,1}K_1 & \cdots & -H_{v,1}K_v \\ \vdots & \ddots & \vdots \\ -H_{1,v}K_1 & \cdots & I - H_{v,v}K_v \end{bmatrix} \text{ is invertible.}$$

Then, for all $d_1, d_2, d_3, d_4 \in \mathbb{Z}^+$, *the following two conditions are equivalent.*

$$\begin{aligned} &(i)\, \operatorname{rank} K_{rx1}(z)G_{tx1,rx1}(z, K(z))K_{tx1}(z) \geq d_1 \\ &(ii)\, \operatorname{rank} K_{rx2}(z)G_{tx2,rx2}(z, K(z))K_{tx2}(z) \geq d_2 \\ &(iii)\, \operatorname{rank} K_{rx2}(z)G_{tx1,rx2}(z, K(z))K_{tx1}(z) \leq d_3 \\ &(iv)\, \operatorname{rank} K_{rx1}(z)G_{tx2,rx1}(z, K(z))K_{tx2}(z) \leq d_4 \end{aligned}$$

if and only if

$$\begin{aligned} &(a)\, \operatorname{rank} G_{tx',rx11'}(z, K(z)) \geq d + d_1 \\ &(b)\, \operatorname{rank} G_{tx',rx22'}(z, K(z)) \geq d + d_2 \\ &(c)\, \operatorname{rank} G_{tx',rx12'}(z, K(z)) \leq d + d_3 \\ &(d)\, \operatorname{rank} G_{tx',rx21'}(z, K(z)) \leq d + d_4 \end{aligned}$$

Proof. Similar to Lemma 3.3. □

Remark 1. The linearized problem of this theorem is essentially the same as that of broadcast problems. Compared with Theorem 8.2, the only difference is that B_{tx1} and B_{tx2} of $G^{lin}_{uni}(z, K)$ are different in multiple-unicast problems while they are the same in broadcast problems.

Remark 2. Like the broadcast problem, to design a two-unicast scheme which communicates a rate r_1 message to receiver 1 and a rate r_2 message to receiver 2, we have to choose $d_1 = r_1$, $d_2 = r_2$, $d_3 = 0$, $d_4 = 0$. The linearized network of Figure 8.7 can be understood as a two-receiver and two-eavesdropper secrecy problem.

8.2 Jordan Form Externalization Example

In this section, we show how the Jordan form externalization of the implicit communication works by working out an explicit example. Let

$$A = \begin{bmatrix} \begin{matrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{matrix} & 0 & 0 \\ 0 & \begin{matrix} \lambda & 1 \\ 0 & \lambda \end{matrix} & 0 \\ 0 & 0 & \lambda' \end{bmatrix}$$

$$C_i = \begin{bmatrix} C_{i,1} & C_{i,2} & C_{i,3} & C_{i,4} & C_{i,5} & C_{i,6} \end{bmatrix}$$

$$B_i = \begin{bmatrix} B_{i,1} \\ B_{i,2} \\ B_{i,3} \\ B_{i,4} \\ B_{i,5} \\ B_{i,6} \end{bmatrix}$$

where $\lambda \neq \lambda'$, $B_{i,j}$ are row vectors, $C_{i,j}$ are column vectors. We will externalize at the frequency $z = \lambda$.

As mentioned in Section 3.5.2, we will move the third and fifth rows and the first and fourth columns of $\lambda I - A$ to the left-top of the matrix. For this, we will define the permutation matrices $P_{L,\lambda}$ and $P_{R,\lambda}$.

The definitions of Section 3.5.2 is given as follows:

$$\kappa_{L,\lambda}(0) = 0, \kappa_{L,\lambda}(1) = 0, \kappa_{L,\lambda}(2) = 0, \kappa_{L,\lambda}(3) = 1, \kappa_{L,\lambda}(4) = 1, \kappa_{L,\lambda}(5) = 2, \kappa_{L,\lambda}(6) = 2$$

$$\kappa_{R,\lambda}(0) = 0, \kappa_{R,\lambda}(1) = 1, \kappa_{R,\lambda}(2) = 1, \kappa_{R,\lambda}(3) = 1, \kappa_{R,\lambda}(4) = 2, \kappa_{R,\lambda}(5) = 2, \kappa_{R,\lambda}(6) = 2$$

$$m_\lambda = 2$$

$$\iota_{L,\lambda}(0) = 0, \iota_{L,\lambda}(1) = 3, \iota_{L,\lambda}(2) = 5$$

$$\iota_{R,\lambda}(0) = 0, \iota_{R,\lambda}(1) = 1, \iota_{L,\lambda}(2) = 4$$

$$\pi_{L,\lambda}(1) = 3, \pi_{L,\lambda}(2) = 4, \pi_{L,\lambda}(3) = 1, \pi_{L,\lambda}(4) = 5, \pi_{L,\lambda}(5) = 2, \pi_{L,\lambda}(6) = 6$$

$$\pi_{R,\lambda}(1) = 1, \pi_{R,\lambda}(2) = 3, \pi_{R,\lambda}(3) = 4, \pi_{R,\lambda}(4) = 2, \pi_{R,\lambda}(5) = 5, \pi_{R,\lambda}(6) = 6$$

$$P_{L,\lambda} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, P_{R,\lambda} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

By multiplying $P_{L,\lambda}^T$ and $P_{R,\lambda}$ to the left and right side of $(zI - A)$, we get the following:

$$P_{L,\lambda}^T (zI - A) P_{R,\lambda} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}^T \begin{bmatrix} z-\lambda & -1 & 0 & & & \\ 0 & z-\lambda & -1 & 0 & & 0 \\ 0 & 0 & z-\lambda & & & \\ & & & z-\lambda & -1 & \\ & 0 & & 0 & z-\lambda & 0 \\ & 0 & & & 0 & z-\lambda' \end{bmatrix}$$

$$\cdot \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & z-\lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & z-\lambda & 0 \\ z-\lambda & -1 & 0 & 0 & 0 & 0 \\ 0 & z-\lambda & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & z-\lambda & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & z-\lambda' \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 & z-\lambda & 0 & 0 \\ 0 & 0 & 0 & 0 & z-\lambda & 0 \\ z-\lambda & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & z-\lambda & -1 & 0 & 0 \\ 0 & z-\lambda & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & z-\lambda' \end{bmatrix}$$

Here, we can notice that the 2×2 left-top sub-matrix is a zero matrix. Furthermore, $P_{L,\lambda}^T (\lambda I - A) P_{R,\lambda}$ is a diagonal matrix.

$A_{\lambda,1,1}(z), A_{\lambda,1,2}(z), A_{\lambda,2,1}(z), A_{\lambda,2,2}(z)$ are defined as

$$A_{\lambda,1,1}(z) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, A_{\lambda,1,2}(z) = \begin{bmatrix} 0 & z-\lambda & 0 & 0 \\ 0 & 0 & z-\lambda & 0 \end{bmatrix}$$

$$A_{\lambda,2,1}(z) = \begin{bmatrix} z-\lambda & 0 \\ 0 & 0 \\ 0 & z-\lambda \\ 0 & 0 \end{bmatrix}, A_{\lambda,2,2}(z) = \begin{bmatrix} -1 & 0 & 0 & 0 \\ z-\lambda & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & z-\lambda' \end{bmatrix}.$$

We also multiply $P_{R,\lambda}$ and $P_{L,\lambda}$ to C_i and B_i respectively.

$$C_i P_{R,\lambda} = \begin{bmatrix} C_{i,1} & C_{i,4} & C_{i,2} & C_{i,3} & C_{i,5} & C_{i,6} \end{bmatrix}$$

$$P_{L,\lambda}^T B_i = \begin{bmatrix} B_{i,3} \\ B_{i,5} \\ B_{i,1} \\ B_{i,2} \\ B_{i,4} \\ B_{i,6} \end{bmatrix}$$

Therefore, $C_{i,\lambda,1}, C_{i,\lambda,2}, B_{i,\lambda,1}, B_{i,\lambda,2}$ are defined as follows.

$$C_{i,\lambda,1} = \begin{bmatrix} C_{i,1} & C_{i,4} \end{bmatrix}, C_{i,\lambda,2} = \begin{bmatrix} C_{i,2} & C_{i,3} & C_{i,5} & C_{i,6} \end{bmatrix}$$

$$B_{i,\lambda,1} = \begin{bmatrix} B_{i,3} \\ B_{i,5} \end{bmatrix}, B_{i,\lambda,2} = \begin{bmatrix} B_{i,1} \\ B_{i,2} \\ B_{i,4} \\ B_{i,6} \end{bmatrix}$$

We also introduce auxiliary inputs and outputs which access each Jordan block. For this, we define C_λ and B_λ as follows.

$$C_\lambda = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}, B_\lambda = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

With these definitions, we can construct the network $\mathcal{N}_{jd.\lambda}$. The channel matrices of

$\mathcal{N}_{jd,\lambda}(\lambda)$ are given as follows:

$$
\begin{aligned}
&H_{tx,rx}(\lambda) = 0 \\
&H_{tx,i}(\lambda) = \begin{bmatrix} C_{i,1} & C_{i,4} \end{bmatrix} \\
&H_{i,rx}(\lambda) = \begin{bmatrix} B_{i,3} \\ B_{i,5} \end{bmatrix} \\
&H_{i,j}(\lambda) = \begin{bmatrix} C_{i,2} & C_{i,3} & C_{i,5} & C_{i,6} \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & \lambda - \lambda' \end{bmatrix}^{-1} \begin{bmatrix} B_{i,1} \\ B_{i,2} \\ B_{i,4} \\ B_{i,6} \end{bmatrix}
\end{aligned}
$$

8.3 Externalization of Implicit Communication in Proper Systems

In this section, we extend the discussion of Section 3.5 to proper systems. The extension of fixed modes to proper systems can be found in [23]. Formally, a proper decentralized linear system, $\mathcal{L}(A, B_i, C_i, D_{ij})$, is defined as follows:

$$
\begin{aligned}
x[n+1] &= Ax[n] + \sum_{i=1}^{v} B_i u_i[n] \\
y_i[n] &= C_i x[n] + \sum_{j=1}^{v} D_{ij} u_j[n]
\end{aligned}
$$

Unlike strictly proper systems, the observations $y_i[n]$ depend not only on the states but also the control inputs $u_i[n]$. Then, the definition of fixed modes can be extended to proper decentralized systems as follows.

Definition 8.1. *[23, Definition 2] λ is called a fixed mode of $\mathcal{L}(A, B_i, C_i, D_{ij})$ if*

$$
\lambda \in \bigcap_{(K_1,\cdots,K_i)\in\mathcal{K}} \sigma(A + \begin{bmatrix} B_1K_1 & \cdots & B_vK_v \end{bmatrix} \left(I - \begin{bmatrix} D_{11}K_1 & \cdots & D_{1v}K_v \\ \vdots & \ddots & \vdots \\ D_{v1}K_1 & \cdots & D_{vv}K_v \end{bmatrix} \right)^{-1} \begin{bmatrix} C_1 \\ \vdots \\ C_v \end{bmatrix})
$$

where $\sigma(\cdot)$ is the set of the eigenvalues of the matrix and $\mathcal{K} = \{(K_1, \cdots, K_v) : K_i \in \mathbb{C}^{q_i \times r_i}, I - \begin{bmatrix} D_{11}K_1 & \cdots & D_{1v}K_v \\ \vdots & \ddots & \vdots \\ D_{v1}K_1 & \cdots & D_{vv}K_v \end{bmatrix}$ is invertible$\}$.

As before, the stabilizability condition is charaterized by the fixed modes of the system.

Theorem 8.4. *[23, Theorem 3] $\mathcal{L}(A, B_i, C_i, D_{ij})$ is stabilizable if and only if all of its fixed modes are within the unit circle.*

Then, we can externalize information flows to stabilize the proper system as before.

8.4 Canonical Externalization I

We will introduce the gain K_i to the ith controller, and the auxiliary input $u[n]$ and output $y[n]$ (which can access all states and observations, $x[n], y_1[n], \cdots, y_v[n]$) to the system. Then, the system equation can be written as follows:

$$\begin{bmatrix} x[n+1] \\ y_1[n] \\ \vdots \\ y_v[n] \end{bmatrix} = \begin{bmatrix} A & B_1K_1 & \cdots & B_vK_v \\ C_1 & D_{11}K_1 & \cdots & D_{1v}K_v \\ \vdots & \vdots & \ddots & \vdots \\ C_v & D_{v1}K_1 & \cdots & D_{vv}K_v \end{bmatrix} \begin{bmatrix} x[n] \\ y_1[n] \\ \vdots \\ y_v[n] \end{bmatrix} + u[n]$$

$$y[n] = \begin{bmatrix} x[n] \\ y_1[n] \\ \vdots \\ y_v[n] \end{bmatrix}$$

Then, the transfer function from $y(z)$ to $u(z)$, $G_{cnI}(z, K)$, is given as follows.

$$\begin{aligned} G_{cnI}(z, K_i) &= \begin{bmatrix} zI & 0 & \cdots & 0 \\ 0 & I & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & I \end{bmatrix} - \begin{bmatrix} A & B_1K_1 & \cdots & B_vK_v \\ C_1 & D_{11}K_1 & \cdots & D_{1v}K_v \\ \vdots & \vdots & \ddots & \vdots \\ C_v & D_{v1}K_1 & \cdots & D_{vv}K_v \end{bmatrix} \\ &= \underbrace{\begin{bmatrix} zI-A & 0 & \cdots & 0 \\ -C_1 & I & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -C_v & 0 & \cdots & I \end{bmatrix}}_{:=A_{cnI}(z)} + \underbrace{\begin{bmatrix} B_1 \\ D_{11} \\ \vdots \\ D_{v1} \end{bmatrix}}_{:=B_{cnI,1}} K_1 \underbrace{\begin{bmatrix} 0 & -I & \cdots & 0 \end{bmatrix}}_{:=C_{cnI,1}} + \cdots + \underbrace{\begin{bmatrix} B_v \\ D_{1v} \\ \vdots \\ D_{vv} \end{bmatrix}}_{:=B_{cnI,v}} K_v \underbrace{\begin{bmatrix} 0 & 0 & \cdots & -I \end{bmatrix}}_{:=C_{cnI,v}} \end{aligned}$$

By Lemma 3.5, the standard network, $\mathcal{N}_s(A_{cnI}(z); B_{cnI,i}, 0; C_{cnI,i}, 0; 0, 0; 0, 0)$, has $G_{cnI}(z, K)$ as a transfer function. Denote this network as $\mathcal{N}_{cnI}(z)$. Then, we can prove the similar theorem as before.

Theorem 8.5. *Given the above definitions, the following statements are equivalent.*

(1) λ *is a fixed mode of the decentralized linear system* $\mathcal{L}(A, B_i, C_i, D_{ij})$

(2) $\mathrm{rank}(G_{cnI}(\lambda, K)) < \dim(A_{cnI})$

(3) (*transfer matrix rank of the LTI network* $\mathcal{N}_{cnI}(\lambda)$) $< \dim(A_{cnI})$

(4) (*mincut rank of the LTI network* $\mathcal{N}_{cnI}(\lambda)$) $< \dim(A_{cnI})$

(5) $\min_{V \subseteq \{1,\cdots,v\}} \mathrm{rank} \begin{bmatrix} A_{cnI}(\lambda) & B_{cnI,V} \\ C_{cnI,V^c} & 0 \end{bmatrix} < \dim(A_{cnI})$

Proof. Similar to Theorem 3.7. After we define $G_{cnI}(z, K)$ as above, the D_{ij} are just a part of $B_{cnI,i}$. □

8.5 Canonical Externalization II

Like the discussion of Section 3.5, we only need the auxiliary input and output to be connected to the unstable states. Thus, we can reduce the dimension of the auxiliary input and output by allowing them only to access the state $x[n]$. Now, the system equation is given as follows:

$$\begin{bmatrix} x[n+1] \\ y_1[n] \\ \vdots \\ y_v[n] \end{bmatrix} = \begin{bmatrix} A & B_1K_1 & \cdots & B_vK_v \\ C_1 & D_{11}K_1 & \cdots & D_{1v}K_v \\ \vdots & \vdots & \ddots & \vdots \\ C_v & D_{v1}K_1 & \cdots & D_{vv}K_v \end{bmatrix} \begin{bmatrix} x[n] \\ y_1[n] \\ \vdots \\ y_v[n] \end{bmatrix} + \begin{bmatrix} I \\ 0 \\ \vdots \\ 0 \end{bmatrix} u[n]$$

$$y[n] = \begin{bmatrix} I & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} x[n] \\ y_1[n] \\ \vdots \\ y_v[n] \end{bmatrix}$$

The transfer function from $u(z)$ to $y(z)$ is the following.

$$y(z) = \begin{bmatrix} I & 0 & \cdots & 0 \end{bmatrix} \left(\begin{bmatrix} zI & 0 & \cdots & 0 \\ 0 & I & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & I \end{bmatrix} - \begin{bmatrix} A & B_1K_1 & \cdots & B_vK_v \\ C_1 & D_{11}K_1 & \cdots & D_{1v}K_v \\ \vdots & \vdots & \ddots & \vdots \\ C_v & D_{v1}K_1 & \cdots & D_{vv}K_v \end{bmatrix} \right)^{-1} \begin{bmatrix} I \\ 0 \\ \vdots \\ 0 \end{bmatrix} u(z)$$

By Lemma 3.6, the transfer function from $y(z)$ to $u(z)$, $G_{cnII}(z,K)$, is given as follows:

$$G_{cnII}(z,K) = (zI-A) - \begin{bmatrix} -B_1K_1 & \cdots & -B_vK_v \end{bmatrix} \left(I - \begin{bmatrix} D_{11}K_1 & \cdots & D_{1v}K_v \\ \vdots & \ddots & \vdots \\ D_{v1}K_1 & \cdots & D_{vv}K_v \end{bmatrix} \right)^{-1} \begin{bmatrix} -C_1 \\ \vdots \\ -C_v \end{bmatrix}$$

$$= \underbrace{(zI-A)}_{:=A_{cnII}(z)} + (\underbrace{-B_1}_{:=B_{cnII,1}} K_1 \underbrace{\begin{bmatrix} I & \cdots & 0 \end{bmatrix}}_{:=C'_{cnII,1}} - \cdots \underbrace{-B_v}_{:=B_{cnII,v}} K_v \underbrace{\begin{bmatrix} 0 & \cdots & I \end{bmatrix}}_{:=C'_{cnII,v}})$$

$$\cdot \left(\underbrace{I}_{:=S^{-1}_{cnII}} - \left(\underbrace{\begin{bmatrix} D_{11} \\ \vdots \\ D_{v1} \end{bmatrix}}_{:=B'_{cnII,1}} K_1 \begin{bmatrix} I & \cdots & 0 \end{bmatrix} + \cdots + \underbrace{\begin{bmatrix} D_{1v} \\ \vdots \\ D_{vv} \end{bmatrix}}_{:=B'_{cnII,v}} K_v \begin{bmatrix} 0 & \cdots & I \end{bmatrix} \right) \right)^{-1} \underbrace{\begin{bmatrix} C_1 \\ \vdots \\ C_v \end{bmatrix}}_{:=D'_{cnII}}$$

Then, by Lemma 3.5, we can see that $G_{cnII}(z,K)$ is the transfer function of the standard network $\mathcal{N}_s(A_{cnII}(z); B_{cnII,i}, B'_{cnII,i}; 0, C'_{cnII,i}; 0, D'_{cnII}; S_{cnII}, 0)$. Denote this network as $\mathcal{N}_{cnII}(z)$. Furthermore, by Lemma 3.5 the channel between the nodes and the channel for the cut $V = \{tx, i_1, \cdots, i_k\}$ is given as follows:

$$H_{tx,rx}(z) = zI - A$$
$$H_{tx,i} = C_i$$
$$H_{i,rx} = -B_i$$
$$H_{i,j} = D_{ji}$$
$$H_{V,V^c}(z) = \begin{bmatrix} zI-A & -B_{i_1} & \cdots & -B_{i_k} \\ C_{i_{k+1}} & D_{i_{k+1},i_1} & \cdots & D_{i_{k+1},i_k} \\ \vdots & \vdots & \ddots & \vdots \\ C_{i_v} & D_{i_v,i_1} & \cdots & D_{i_v,i_k} \end{bmatrix}$$

Then, we can give the capacity-stabilizability equivalence theorem as before.

Theorem 8.6. *Given the above definitions, the following statements are equivalent.*

(1) λ *is a fixed mode of the decentralized linear system* $\mathcal{L}(A, B_i, C_i, D_{ij})$

(2) $\text{rank}(G_{cnII}(\lambda, K)) < \dim(A)$

(3) (*transfer matrix rank of the LTI network* $\mathcal{N}_{cnII}(\lambda)$) $< \dim(A)$

(4) (*mincut rank of the LTI network* $\mathcal{N}_{cnII}(\lambda)$) $< \dim(A)$

(5) $\min_{V \subseteq \{1,\cdots,v\}} \text{rank} \begin{bmatrix} \lambda I - A & -B_V \\ C_{V^c} & D_{V^c,V} \end{bmatrix} < \dim(A)$

Proof. Similar to Theorem 3.7. □

Here, it has to be mentioned that the equivalence of (1) and (5) was already shown in [23].

8.6 Jordan Form Externalization

Like Section 3.5.2, we can minimize the dimension of the auxiliary input and output by using the Jordan form. Without loss of generality, we assume that A is in Jordan form and use the same notations of Section 3.5.2. Then, the system equation with the auxiliary input $u_\lambda[n]$ and output $y_\lambda[n]$ is given as follows:

$$\begin{bmatrix} x[n+1] \\ y_1[n] \\ \vdots \\ y_v[n] \end{bmatrix} = \begin{bmatrix} A & B_1K_1 & \cdots & B_vK_v \\ C_1 & D_{11}K_1 & \cdots & D_{1v}K_v \\ \vdots & \vdots & \ddots & \vdots \\ C_v & D_{v1}K_1 & \cdots & D_{vv}K_v \end{bmatrix} \begin{bmatrix} x[n] \\ y_1[n] \\ \vdots \\ y_v[n] \end{bmatrix} + \underbrace{\begin{bmatrix} C_\lambda \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{:=C'_\lambda} u_\lambda[n]$$

$$y_\lambda[n] = \underbrace{\begin{bmatrix} B_\lambda & 0 & \cdots & 0 \end{bmatrix}}_{:=B'_\lambda} \begin{bmatrix} x[n] \\ y_1[n] \\ \vdots \\ y_v[n] \end{bmatrix}$$

We also expand the dimension of the permutation matrices $P_{L,\lambda}$ and $P_{R,\lambda}$.

$$P'_{L,\lambda} := \begin{bmatrix} P_{L,\lambda} & 0 & \cdots & 0 \\ 0 & I & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & I \end{bmatrix}$$

$$P'_{R,\lambda} := \begin{bmatrix} P_{R,\lambda} & 0 & \cdots & 0 \\ 0 & I & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & I \end{bmatrix}$$

The transfer function from $u_\lambda(z)$ to $y_\lambda(z)$ is the following.

$$y_\lambda(z) = C'_\lambda(\begin{bmatrix} zI & 0 & \cdots & 0 \\ 0 & I & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & I \end{bmatrix} - \begin{bmatrix} A & B_1K_1 & \cdots & B_vK_v \\ C_1 & D_{11}K_1 & \cdots & D_{1v}K_v \\ \vdots & \vdots & \ddots & \vdots \\ C_v & D_{v1}K_v & \cdots & D_{vv}K_v \end{bmatrix})^{-1} B'_\lambda u_\lambda(z)$$

$$= C'_\lambda(P'_{L,\lambda}{P'_{L,\lambda}}^T(\begin{bmatrix} zI-A & -B_1K_1 & \cdots & -B_vK_v \\ -C_1 & I-D_{11}K_1 & \cdots & -D_{1v}K_v \\ \vdots & \vdots & \ddots & \vdots \\ -C_v & -D_{v1}K_v & \cdots & I-D_{vv}K_v \end{bmatrix})P'_{R,\lambda}{P'_{R,\lambda}}^T)^{-1} B'_\lambda u_\lambda(z)$$

$$= C'_\lambda P'_{R,\lambda}({P'_{L,\lambda}}^T(\begin{bmatrix} zI-A & -B_1K_1 & \cdots & -B_vK_v \\ -C_1 & I-D_{11}K_1 & \cdots & -D_{1v}K_v \\ \vdots & \vdots & \ddots & \vdots \\ -C_v & -D_{v1}K_v & \cdots & I-D_{vv}K_v \end{bmatrix})P'_{R,\lambda})^{-1}{P'_{L,\lambda}}^T B'_\lambda u_\lambda(z)$$

$$= C'_\lambda P'_{R,\lambda}\begin{bmatrix} P^T_{L,\lambda}(zI-A)P_{R,\lambda} & -P^T_{L,\lambda}B_1K_1 & \cdots & -P^T_{L,\lambda}B_vK_v \\ -C_1P_{R,\lambda} & I-D_{11}K_1 & \cdots & -D_{1v}K_v \\ \vdots & \vdots & \ddots & \vdots \\ -C_vP_{R,\lambda} & -D_{v1}K_v & \cdots & I-D_{vv}K_v \end{bmatrix}^{-1}{P'_{L,\lambda}}^T B'_\lambda u_\lambda(z)$$

$$= \begin{bmatrix} I & 0 & 0 & \cdots & 0 \end{bmatrix}\begin{bmatrix} A_{\lambda,1,1}(z) & A_{\lambda,1,2}(z) & -B_{1,\lambda,1}K_1 & \cdots & -B_{v,\lambda,1}K_v \\ A_{\lambda,2,1}(z) & A_{\lambda,2,2}(z) & -B_{1,\lambda,2}K_1 & \cdots & -B_{v,\lambda,2}K_v \\ -C_{1,\lambda,1} & -C_{1,\lambda,2} & I-D_{11}K_1 & \cdots & -D_{1v}K_v \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -C_{v,\lambda,1} & -C_{v,\lambda,2} & -D_{v1}K_v & \cdots & I-D_{vv}K_v \end{bmatrix}^{-1}\begin{bmatrix} I \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} u_\lambda(z)$$

By Lemma 3.6, the transfer matrix $G_{jd}(z)$ from $y_\lambda(z)$ to $u_\lambda(z)$ is given as

$$G_{jd}(z) = A_{\lambda,1,1}(z) - \begin{bmatrix} A_{\lambda,1,2}(z) & -B_{1,\lambda,1}K_1 & \cdots & -B_{v,\lambda,1}K_v \end{bmatrix}$$

$$\cdot \begin{bmatrix} A_{\lambda,2,2}(z) & -B_{1,\lambda,2}K_1 & \cdots & -B_{v,\lambda,2}K_v \\ -C_{1,\lambda,2} & I - D_{11}K_1 & \cdots & -D_{1v}K_v \\ \vdots & \vdots & \ddots & \vdots \\ -C_{v,\lambda,2} & -D_{v1}K_1 & \cdots & I - D_{vv}K_v \end{bmatrix}^{-1} \begin{bmatrix} A_{\lambda,2,1}(z) \\ -C_{1,\lambda,1} \\ \vdots \\ -C_{v,\lambda,1} \end{bmatrix}$$

$$= \underbrace{A_{\lambda,1,1}(z)}_{:=A_{jd}(z)} + (\underbrace{\begin{bmatrix} A_{\lambda,1,2}(z) & 0 & \cdots & 0 \end{bmatrix}}_{:=D_{jd}(z)} \underbrace{-B_{1,\lambda,1}}_{:=B_{jd,1}} K_1 \underbrace{\begin{bmatrix} 0 & I & \cdots & 0 \end{bmatrix}}_{:=C'_{jd,1}} - \cdots \underbrace{-B_{v,\lambda,1}}_{:=B_{jd,v}} K_v \underbrace{\begin{bmatrix} 0 & 0 & \cdots & I \end{bmatrix}}_{:=C'_{jd,v}})$$

$$\cdot (\underbrace{I}_{:=S_{jd}^{-1}} - (\underbrace{\begin{bmatrix} I - A_{\lambda,2,2}(z) & 0 & \cdots & 0 \\ C_{1,\lambda,2} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ C_{v,\lambda,2} & 0 & \cdots & 0 \end{bmatrix}}_{:=S'_{jd}(z)} + \underbrace{\begin{bmatrix} B_{1,\lambda,2} \\ D_{11} \\ \vdots \\ D_{v1} \end{bmatrix}}_{:=B'_{jd,1}} K_1 \begin{bmatrix} 0 & I & \cdots & 0 \end{bmatrix} + \cdots + \underbrace{\begin{bmatrix} B_{v,\lambda,2} \\ D_{1v} \\ \vdots \\ D_{vv} \end{bmatrix}}_{:=B'_{jd,v}} K_v \begin{bmatrix} 0 & 0 & \cdots & I \end{bmatrix}))^{-1}$$

$$\cdot \underbrace{\begin{bmatrix} -A_{\lambda,2,1}(z) \\ C_{1,\lambda,1} \\ \vdots \\ C_{v,\lambda,1} \end{bmatrix}}_{:=D'_{jd}(z)}$$

Then, we can easily check that $G_{jd}(z)$ is the transfer function of the standard network

$$\mathcal{N}_s(A_{jd}(z); B_{jd,i}, B'_{jd,i}; 0, C'_{jd,i}; D_{jd}(z), D'_{jd}(z); S_{jd}, S'_{jd}(z)).$$

Moreover, we have

$$(S_{jd}^{-1} - S'_{jd})^{-1} = \begin{bmatrix} A_{\lambda,2,2}(z) & 0 & \cdots & 0 \\ -C_{1,\lambda,2}(z) & I & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -C_{v,\lambda,2}(z) & 0 & \cdots & I \end{bmatrix}^{-1} = \begin{bmatrix} A_{\lambda,2,2}(z)^{-1} & 0 & \cdots & 0 \\ C_{1,\lambda,2}A_{\lambda,2,2}(z)^{-1} & I & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ C_{v,\lambda,2}A_{\lambda,2,2}(z)^{-1} & 0 & \cdots & I \end{bmatrix}.$$

Thus, by Lemma 3.5 the channel matrix between the nodes and the channel matrix for the cut $V = \{tx, i_1, \cdots, i_k\}$ is given as follows:

$$H_{tx,rx}(\lambda) = 0$$

$$H_{tx,i}(\lambda) = C_{i,\lambda,1}$$

$$H_{i,rx}(\lambda) = -B_{i,\lambda,1}$$

$$H_{i,j}(\lambda) = C_{j,\lambda,2}A_{\lambda,2,2}(\lambda)^{-1}B_{i,\lambda,2} + D_{ji}$$

$$H_{V,V^c}(\lambda) := \begin{bmatrix} 0 & -B_{i_1,\lambda,1} & \cdots & -B_{i_k,\lambda,1} \\ C_{i_{k+1},\lambda,1} & C_{i_{k+1},\lambda,2}A_{\lambda,2,2}(\lambda)^{-1}B_{i_1,\lambda,2} + D_{i_{k+1}i_1} & \cdots & C_{i_{k+1},\lambda,2}A_{\lambda,2,2}(\lambda)^{-1}B_{i_k,\lambda,2} + D_{i_{k+1}i_k} \\ \vdots & \vdots & \ddots & \vdots \\ C_{i_v,\lambda,1} & C_{i_v,\lambda,2}A_{\lambda,2,2}(\lambda)^{-1}B_{i_1,\lambda,2} + D_{i_v i_1} & \cdots & C_{i_v,\lambda,2}A_{\lambda,2,2}(\lambda)^{-1}B_{i_k,\lambda,2} + D_{i_v i_k} \end{bmatrix}$$

Then, we can write a similar theorem as before.

Theorem 8.7. *Given the above definitions, the following statements are equivalent.*

(1) λ *is the fixed mode of the decentralized linear system* $\mathcal{L}(A, B_i, C_i, D_{ij})$

(2) $\text{rank}(G_{jd}(\lambda, K)) < m_\lambda$

(3) (*transfer matrix rank of the LTI network* $\mathcal{N}_{jd}(\lambda)$) $< m_\lambda$

(4) (*mincut rank of the LTI network* $\mathcal{N}_{jd}(\lambda)$) $< m_\lambda$

(5) $$\min_{V \subseteq \{1,\cdots,v\}} \text{rank} \begin{bmatrix} 0 & -B_{V,\lambda,1} \\ C_{V^c,\lambda,1} & C_{V^c,\lambda,2}A_{\lambda,2,2}(\lambda)^{-1}B_{V,\lambda,2} + D_{V^c,V} \end{bmatrix} < m_\lambda$$

Proof. Similar to Theorem 3.8. Compared to Theorem 3.8, $D_{V^c,V}$ is just added to $C_{V^c,\lambda,2}A_{\lambda,2,2}(\lambda)^{-1}B_{V,\lambda,2}$. □

8.7 Realization of Closed LTI Network

In this section, we will discuss how to realize the problem of Figure 3.17 in a decentralized linear system form. First, we can notice that the system of Figure 3.17 can be thought as a special case of the closed LTI network of Figure 8.8. We can put p of Figure 8.8 as $v + 2$, and consider the relay i of Figure 3.17 as the node i of Figure 8.8, the observer as the node $v + 1$, and the controller as the node $v + 2$. Then, by connecting the node $v + 1$ with the node $v + 2$ with $H_{(v+2)(v+1)}(z)$ which is equivalent to the plant of Figure 3.17, the two problems are equivalent. Therefore, we can focus on the realization of the closed LTI network of Fig. 8.8.

As we can see in Figure 3.17, for $1 \leq i, j \leq p$ the input of node i is connected to the output of node j by the channel $H_{ij}(z)$. When $i = j$, it corresponds to a self-loop. In other words,

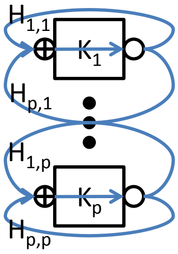


Figure 8.8: General Closed LTI Network

$y_j(z) = H_{ij}(z)u_i(z)$ where $u_i(z)$ is the input of the node i and $y_j(z)$ is the output of the node j. Since this relationship can be considered as a centralized input-output system, it can be realized by the usual realization method shown in [17, chapter 7]. Let's say the resulting linear system is given as follows:

$$x_{ij}[n+1] = A_{ij}x_{ij}[n] + B_{ij}u_i[n]$$
$$y_j[n] = C_{ij}x_{ij}[n] + D_{ij}u_i[n]$$

Let the dimension of $u_i[n]$ be q_i, the dimension of $y_i[n]$ be r_i and the dimension of $x_{ij}[n]$ be m_{ij}. Then, the dimensions of the other matrices are uniquely determined. When there is no connection between the nodes, simply m_{ij} becomes 0.

The main idea for the realization of a closed LTI network is to augment the states $x_{ij}[n]$. Denote $x[n]$, A, B_i and C_i as follows:

$$x[n] := \begin{bmatrix} x_{11}[n+1] \\ \vdots \\ x_{1p}[n+1] \\ x_{21}[n+1] \\ \vdots \\ x_{pp}[n+1] \end{bmatrix}$$

$$A := diag(A_{11}, \cdots, A_{1p}, A_{21}, \cdots, A_{pp})$$

$$B_i := \begin{bmatrix} 0_{(\sum_{1\le j<i}\sum_{1\le k\le p} m_{jk})\times q_i} \\ B_{i1} \\ \vdots \\ B_{ip} \\ 0_{(\sum_{i<j\le p}\sum_{1\le k\le p} m_{jk})\times q_i} \end{bmatrix}$$

$$C'_{ij} := \begin{bmatrix} 0_{r_j\times\sum_{1\le k<j} m_{ik}} & C_{ij} & 0_{r_j\times\sum_{j<k\le p} m_{ik}} \end{bmatrix}$$

$$C_i := \begin{bmatrix} C'_{1i} & \cdots & C'_{pi} \end{bmatrix}.$$

Then, we can easily check that the decentralized linear system

$$x[n+1] = Ax[n] + \sum_i B_i u_i[n]$$
$$y_i[n] = C_i x[n] + \sum_{i,j} D_{ij} u_i[n]$$

is the realization of the closed LTI network of Fig. 8.8.

Chapter 9

Appendix for Chapter 4

9.1 Proof of Corollary 4.1 of Page 213

Proof of (c):

Let's first consider when $\max(1, a^2\sigma_{v1}^2) = a^2\sigma_{v1}^2$. Since $a^2\sigma_{v1}^2 \geq 1$, there exists $k_1 \geq 2$ such that

$$a^{2(k_1-2)} \leq a^2\sigma_{v1}^2 < a^{2(k_1-1)},$$

and we choose such a k_1 as k_1 in Lemma 4.13. Then, by (4.32) of Lemma 4.13 we have

$$\begin{aligned}
& D_{L,3}(\widetilde{P_1}, \widetilde{P_2}; k_1) \\
& \geq \frac{a^{2(k_1-1)}\sigma_{v1}^2}{(1+\frac{\sigma_{v1}^2}{\sigma_{v2}^2})(\frac{a^{2(k_1-2)}}{1-a^{-2}}) + \sigma_{v1}^2} \\
& \overset{(A)}{\geq} \frac{a^{2(k_1-1)}\sigma_{v1}^2}{\frac{2}{1-2.5^{-2}}a^{2(k_1-2)} + \sigma_{v1}^2} \\
& = \frac{a^2\sigma_{v1}^2}{\frac{2}{1-2.5^{-2}} + \frac{\sigma_{v1}^2}{a^{2(k_1-2)}}} \\
& \overset{(B)}{\geq} \frac{a^2\sigma_{v1}^2}{\frac{2}{1-2.5^{-2}} + 1} \\
& \geq 0.295775...a^2\sigma_{v1}^2 \\
& \geq 0.295a^2\sigma_{v1}^2. \qquad (9.1)
\end{aligned}$$

(A): $\sigma_{v1}^2 \leq \sigma_{v2}^2$ and $|a| \geq 2.5$.
(B): $a^2\sigma_{v1}^2 < a^{2(k_1-1)}$.

When $\max(1, a^2\sigma_{v1}^2) = 1$, by (4.32) of Lemma 4.13 we have $D_{L,3}(\widetilde{P_1}, \widetilde{P_2}; k_1) \geq 1 \geq 0.295$.

Proof of (b):

In Lemma 4.13, choose k_1 in the same way as (c) and let $k = k_1 + 1$. Inspired by the proof of (c), we can safely choose $\Sigma = 0.295\max(1, a^2\sigma_{v1}^2)$. Then, by Lemma 4.13, we notice that since $k - k_1 - 1 = 0$, the second and third square-root terms in $D_{L,2}(\widetilde{P_1}, \widetilde{P_2}; k_1, k, \Sigma)$ goes away and the bound reduces to

$$D_{L,2}(\widetilde{P_1}, \widetilde{P_2}; k_1, k, \Sigma) \geq \inf_{c_1,c_2} (\sqrt{(a - c_1 - c_2)^2\Sigma + c_1^2\sigma_{v1}^2 + c_2^2\sigma_{v2}^2})_+^2 + 1$$
$$\text{s.t. } (1 - 2.5^{-1})c_1^2(\Sigma + \sigma_{v1}^2) \leq \widetilde{P_1}$$

where c_2 can be chosen arbitrarily.

Here, since we assumed $\widetilde{P_1} \leq \frac{1}{400}a^2\max(1, a^2\sigma_{v1})$, we have

$$(1 - 2.5^{-1})c_1^2(\Sigma + \sigma_{v1}^2) \leq \widetilde{P_1} \leq \frac{1}{400}a^2\max(1, a^2\sigma_{v1}^2)$$
$$(\Rightarrow)c_1^2(0.295\max(1, a^2\sigma_{v1}^2) + \sigma_{v1}^2) \leq \frac{1}{400(1 - 2.5^{-1})}a^2\max(1, a^2\sigma_{v1}^2)$$
$$(\Rightarrow)c_1^2 \leq \frac{a^2\max(1, a^2\sigma_{v1}^2)}{400(1 - 2.5^{-1})(0.295\max(1, a^2\sigma_{v1}^2) + \sigma_{v1}^2)} \leq \frac{a^2}{400(1 - 2.5^{-1}) \cdot 0.295}$$
$$(\Rightarrow)|c_1| \leq 0.118846...|a| \leq 0.119|a|.$$

Therefore,

$$\begin{aligned}
D_{L,2}(\widetilde{P_1}, \widetilde{P_2}; k_1, k, \Sigma) &\geq \inf_{c_1,c_2} (\sqrt{(a - c_1 - c_2)^2 0.295\max(1, a^2\sigma_{v1}^2) + c_1^2\sigma_{v1}^2 + c_2^2\sigma_{v2}^2})_+^2 + 1 \\
&\quad \text{s.t. } |c_1| \leq 0.119|a| \\
&\geq \inf_{c_2}(a - 0.119a - c_2)^2 0.295\max(1, a^2\sigma_{v1}^2) + c_2^2\sigma_{v2}^2 + 1 \\
&\overset{(A)}{\geq} \inf_{c_2}(a - 0.119a - c_2)^2 0.295\sigma_{v2}^2 + c_2^2\sigma_{v2}^2 + 1 \\
&= \inf_{\widetilde{c_2}}(\sqrt{(1 - 0.119 - \widetilde{c_2})^2 0.295 + \widetilde{c_2}^2})^2 a^2\sigma_{v2}^2 + 1 \\
&\overset{(B)}{=} 0.176808...a^2\sigma_{v2}^2 + 1 \\
&\geq 0.176a^2\sigma_{v2}^2 + 1.
\end{aligned}$$

(A): By the assumption $\max(1, a^2\sigma_{v1}^2) \geq \sigma_{v2}^2$.

(B): By the numerical optimization of the quadratic function.

Proof of (a):

(i) When $\max(1, a^2\sigma_{v2}^2) = a^2\sigma_{v2}^2$

In Lemma 4.13, we will choose k_1 in the same way as (c) and k arbitrarily large. As above, we can safely choose $\Sigma = 0.295a^2\sigma_{v1}^2$. Applying the same arguments as (b), (c) to Lemma 4.13 gives

$$\begin{aligned}
D_{L,2}(\widetilde{P_1},\widetilde{P_2};k_1,k,\Sigma) \geq \inf_{c_1,c_2} (&\sqrt{a^{2(k-k_1-1)}((a-c_1-c_2)^2\Sigma + c_1^2\sigma_{v1}^2 + c_2^2\sigma_{v2}^2)} \\
&- \sqrt{\frac{a^{2(k-k_1-2)}(1-(2.5a^{-2})^{k-k_1-1})}{1-2.5a^{-2}} \frac{\widetilde{P_1}}{(1-2.5^{-1})2.5^{-1}}} \\
&- \sqrt{\frac{a^{2(k-k_1-2)}(1-(2.5a^{-2})^{k-k_1-1})}{1-2.5a^{-2}} \frac{\widetilde{P_2}}{(1-2.5^{-1})2.5^{-1}}})_+^2 + 1 \\
&\text{s.t. } (1-2.5^{-1})c_1^2(\Sigma+\sigma_{v1}^2) \leq \widetilde{P_1} \\
&\quad (1-2.5^{-1})c_2^2(\Sigma+\sigma_{v2}^2) \leq \widetilde{P_2} \\
&\geq (\sqrt{a^{2(k-k_1-1)}0.176a^2\sigma_{v2}^2} \\
&- \sqrt{\frac{a^{2(k-k_1-2)}(1-(2.5a^{-2})^{k-k_1-1})}{1-2.5a^{-2}} \frac{\widetilde{P_1}}{(1-2.5^{-1})2.5^{-1}}} \\
&- \sqrt{\frac{a^{2(k-k_1-2)}(1-(2.5a^{-2})^{k-k_1-1})}{1-2.5a^{-2}} \frac{\widetilde{P_2}}{(1-2.5^{-1})2.5^{-1}}})_+^2 + 1. \qquad (9.2)
\end{aligned}$$

Moreover, we have

$$\begin{aligned}
&\frac{a^{2(k-k_1-2)}(1-(2.5a^{-2})^{k-k_1-1})}{1-2.5a^{-2}} \frac{\widetilde{P_1}}{(1-2.5^{-1})2.5^{-1}} \\
&\overset{(A)}{\leq} \frac{a^{2(k-k_1-2)}}{1-2.5^{-1}} \frac{\widetilde{P_1}}{(1-2.5^{-1})2.5^{-1}} \\
&\overset{(B)}{\leq} \frac{a^{2(k-k_1-2)}}{(1-2.5^{-1})^2 2.5^{-1}} \frac{1}{400} a^2 \max(1, a^2\sigma_{v1}^2) \\
&\leq 0.01736111...a^{2(k-k_1)}\sigma_{v2}^2 \\
&\leq 0.0174a^{2(k-k_1)}\sigma_{v2}^2. \qquad (9.3)
\end{aligned}$$

(A): $|a| \geq 2.5$.
(B): We assumed $\widetilde{P_1} \leq \frac{1}{400}a^2\max(1, a^2\sigma_{v1}^2) = \frac{1}{400}a^4\sigma_{v1}^2 \leq a^2\sigma_{v2}^2$.

Likewise, we also have

$$\begin{aligned}
&\frac{a^{2(k-k_1-2)}(1-(2.5a^{-2})^{k-k_1-1})}{1-2.5a^{-2}} \frac{\widetilde{P_2}}{(1-2.5^{-1})2.5^{-1}} \\
&\leq 0.0174a^{2(k-k_1)}\sigma_{v2}^2. \qquad (9.4)
\end{aligned}$$

Therefore, by plugging (9.3) and (9.4) into (9.2), we get

$$\begin{aligned}
D_{L,2}(\widetilde{P_1},\widetilde{P_2};k_1,k,\Sigma) &\geq (\sqrt{0.176} - \sqrt{0.0174} - \sqrt{0.0174})_+^2 a^{2(k-k_1)}\sigma_{v2}^2 + 1 \\
&\geq 0.02a^{2(k-k_1)}\sigma_{v2}^2 + 1.
\end{aligned}$$

Since k can be chosen arbitrarily large and $|a| > 1$, $\lim_{k\to\infty} D_{L,2}(\widetilde{P_1},\widetilde{P_2};k_1,k,\Sigma) = \infty$.

(ii) When $\max(1, a^2\sigma_{v2}^2) = 1$

We will choose k arbitrarily large in (4.33) of Lemma 4.13. Here the assumptions $\widetilde{P_1} \leq \frac{1}{400}a^2\max(1, a^2\sigma_{v1}^2)$ and $\widetilde{P_2} \leq \frac{1}{400}a^2\max(1, a^2\sigma_{v2}^2)$ reduce to $\widetilde{P_1} \leq \frac{a^2}{400}$ and $\widetilde{P_2} \leq \frac{a^2}{400}$.

Therefore, by (4.33) of Lemma 4.13 and $|a| \geq 2.5$, for all k we have

$$\begin{aligned}
D_{L,4}(\widetilde{P_1}, \widetilde{P_2}; k) &\geq (\sqrt{a^{2(k-1)}} - \sqrt{\frac{a^{2(k-2)}}{(1-2.5^{-1})^2}\frac{a^2}{400}} - \sqrt{\frac{a^{2(k-2)}}{(1-2.5^{-1})^2}\frac{a^2}{400}})_+^2 \\
&= (1 - \sqrt{\frac{1}{400(1-2.5^{-1})^2}} - \sqrt{\frac{1}{400(1-2.5^{-1})^2}})_+^2 a^{2(k-1)} \\
&\geq 0.6a^{2(k-1)}
\end{aligned}$$

Since k can be chosen arbitrarily large, $\lim_{k\to\infty} D_{L,4}(\widetilde{P_1}, \widetilde{P_2}; k) = \infty$.

9.2 Proof of Proposition 4.7

By Lemma 4.14, if there exists $c \geq 1$ such that for all $\widetilde{P_1}, \widetilde{P_2} \geq 0$, $D_U(c\widetilde{P_1}, c\widetilde{P_2}) \leq c \cdot D_L(\widetilde{P_1}, \widetilde{P_2})$, then for all $q, r_1, r_2 \geq 0$ we have

$$\frac{\min_{P_1,P_2\geq 0} qD_U(P_1,P_2) + r_1P_1 + r_2P_2}{\min_{\widetilde{P_1},\widetilde{P_2}\geq 0} qD_L(\widetilde{P_1},\widetilde{P_2}) + r_1\widetilde{P_1} + r_2\widetilde{P_2}} \leq c$$

which finishes the proof. Therefore, we will only prove that such c exists.

(i) When $\widetilde{P_1} \leq \frac{1}{400}a^2\max(1, a^2\sigma_{v1}^2)$ and $\widetilde{P_2} \leq \frac{1}{400}a^2\max(1, a^2\sigma_{v2}^2)$

Lower bound: By Corollary 4.1 (a),

$$D_L(\widetilde{P_1}, \widetilde{P_2}) = \infty$$

Therefore, we do not need the corresponding upper bound.

(ii) When $\widetilde{P_1} \leq \frac{1}{400}a^2\max(1, a^2\sigma_{v1}^2)$ and $\widetilde{P_2} \geq \frac{1}{400}a^2\max(1, a^2\sigma_{v2}^2)$

Lower bound: By Corollary 4.1 (b),

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq 0.176a^2\sigma_{v2}^2 + 1.$$

Upper bound: By Lemma 4.15,

$$\begin{aligned}
(D_U(P_1,P_2), P_1, P_2) &\leq (a^2\sigma_{v2}^2 + 1, 0, a^4\sigma_{v2}^2 + a^2\sigma_{v2}^2 + a^2) \\
&\leq (a^2\sigma_{v2}^2 + 1, 0, 3a^2\max(1, a^2\sigma_{v2}^2)).
\end{aligned}$$

Ratio: Thus, c is upper bounded by

$$c \leq \max(\frac{1}{0.176}, \frac{3}{\frac{1}{400}}) = 1200.$$

(iii) When $\widetilde{P_1} \geq \frac{1}{400}a^2\max(1, a^2\sigma_{v1}^2)$

Lower bound: By Corollary 4.1 (c),

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq 0.295\max(1, a^2\sigma_{v1}^2)$$

Upper bound: By Lemma 4.15,

$$\begin{aligned}(D_U(P_1,P_2),P_1,P_2) &\leq (a^2\sigma_{v1}^2+1, a^4\sigma_{v1}^2+a^2\sigma_{v1}^2+a^2,0)\\ &\leq (2\max(1,a^2\sigma_{v1}^2), 3a^2\max(1,a^2\sigma_{v1}^2),0)\end{aligned}$$

Ratio: c is upper bounded by

$$c \leq \max(\frac{2}{0.295}, \frac{3}{\frac{1}{400}}) = 1200.$$

Therefore, by (i), (ii), (iii), the lemma is true and $c \leq 1200$.

9.3 Proof of Corollary 4.2

For simplicity, we will only prove for the case when $\max(1,a^2\sigma_{v1}^2) = a^2\sigma_{v1}^2$. To prove the case when $\max(1,a^2\sigma_{v1}^2)=1$, we can simply repeat the following proofs with parameters $k_1 = 1$ and $\Sigma = 0.295$. We will also abbreviate $D_{L,1}(\widetilde{P_1},\widetilde{P_2};k_1,k_2,k,\sigma'_{v2},\alpha,\Sigma)$ to $D_{L,1}(\widetilde{P_1},\widetilde{P_2})$.

Proof of (a):
Since $a^2\sigma_{v1}^2 \geq 1$, there exists $k_1 \geq 2$ such that

$$a^{2(k_1-2)} \leq a^2\sigma_{v1}^2 < a^{2(k_1-1)},$$

and we will use such a k_1 in Lemma 4.12.

Since the selection of k_1 is the same as in the proof of Corollary 4.1 (c), by (9.1) we can select $\Sigma = 0.295a^2\sigma_{v1}^2$ in Lemma 4.12.

Let's further choose $k_2 = k = k_1+s+1$, $\alpha = 1$ and $\sigma_{v2} = \sigma'_{v2}$ in Lemma 4.12. Then, since (**??**), (**??**), (4.18) disappear, $D_{L,1}$ in Lemma 4.12 reduces to

$$D_{L,1}(\widetilde{P_1},\widetilde{P_2}) \geq (\sqrt{\frac{a^{2(k_2-k_1)}\Sigma}{2^{2I'(\widetilde{P_1})}}} - \sqrt{\frac{a^{2(k_2-k_1-1)}\widetilde{P_1}}{(1-2.5^{-1})^2}})_+^2 + 1. \tag{9.5}$$

We can bound $I''(\widetilde{P_1})$ as

$$\begin{aligned}I''(\widetilde{P_1}) &\overset{(A)}{\leq} \frac{1}{2}\log(1+\frac{1}{(k_2-k_1-1)\sigma_{v2}^2}(\frac{2a^{2(k_2-2-k_1)}}{1-2.5^{-2}}0.295a^2\sigma_{v1}^2 + \frac{2a^{2(k_2-3-k_1)}}{1-2.5^{-2}}\frac{2.5\frac{1}{70}\frac{\sigma_{v2}^2}{a^{2(s-1)}}}{(1-2.5^{-1})^2}))^{k_2-k_1-1}\\
&= \frac{1}{2}\log(1+\frac{1}{(k_2-k_1-1)\sigma_{v2}^2}(\frac{2}{1-2.5^{-2}}0.295a^{2s}\sigma_{v1}^2 + \frac{2}{1-2.5^{-2}}\frac{2.5\frac{1}{70}}{(1-2.5^{-1})^2}a^{-2}\sigma_{v2}^2))^{k_2-k_1-1}\\
&\overset{(B)}{\leq} \frac{1}{2}\log(1+\frac{1}{k_2-k_1-1}(\frac{2}{1-2.5^{-2}}0.295 + \frac{2}{1-2.5^{-2}}\frac{\frac{1}{70}\frac{1}{2.5}}{(1-2.5^{-1})^2}))^{k_2-k_1-1}\\
&= \frac{1}{2}\log(1+\frac{0.7401738...}{k_2-k_1-1})^{k_2-k_1-1}\\
&\leq \frac{1}{2}\log(1+\frac{0.7402}{k_2-k_1-1})^{k_2-k_1-1}\\
&\leq \frac{1}{2}\log(e^{0.7402}).\end{aligned}$$

(A): Assumptions $\widetilde{P_1} \leq \frac{\sigma_{v2}^2}{70a^{2(s-1)}}$ and $|a| \geq 2.5$.
(B): Assumptions $a^{2s}\sigma_{v1}^2 \leq \sigma_{v2}^2$ and $|a| \geq 2.5$.

Likewise, $I'(\widetilde{P_1})$ is upper bounded as

$$\begin{aligned}
I'(\widetilde{P_1}) &\leq \frac{1}{2}\log(e^{0.7402}) + \frac{1}{2}\log(1 + \frac{1}{\sigma_{v2}^2}(2a^{2(k_2-1-k_1)}\Sigma + 2\frac{a^{2(k_2-2-k_1)}\widetilde{P_1}}{(1-2.5a^{-2})(1-2.5^{-1})})) \\
&\overset{(A)}{\leq} \frac{1}{2}\log(e^{0.7402}) + \frac{1}{2}\log(1 + \frac{1}{\sigma_{v2}^2}(2a^{2s}0.295a^2\sigma_{v1}^2 + 2\frac{a^{2(s-1)}\frac{1}{70}\frac{\sigma_{v2}^2}{a^{2(s-1)}}}{(1-2.5^{-1})^2})) \\
&= \frac{1}{2}\log(e^{0.7402}) + \frac{1}{2}\log(\frac{1}{\sigma_{v2}^2}(\sigma_{v2}^2 + 2\times 0.295a^{2(s+1)}\sigma_{v1}^2 + \frac{1}{35(1-2.5^{-1})^2}\sigma_{v2}^2)) \\
&\overset{(B)}{\leq} \frac{1}{2}\log(e^{0.7402}) + \frac{1}{2}\log(\frac{1}{\sigma_{v2}^2}(a^{2(s+1)}\sigma_{v1}^2 + 2\times 0.295a^{2(s+1)}\sigma_{v1}^2 + \frac{1}{35(1-2.5^{-1})^2}a^{2(s+1)}\sigma_{v1}^2)) \\
&= \frac{1}{2}\log(e^{0.7402}) + \frac{1}{2}\log(\frac{a^{2(s+1)}\sigma_{v1}^2}{\sigma_{v2}^2}(1 + 2\times 0.295 + \frac{1}{35(1-2.5^{-1})^2})) \\
&= \frac{1}{2}\log(e^{0.7402}) + \frac{1}{2}\log(\frac{a^{2(s+1)}\sigma_{v1}^2}{\sigma_{v2}^2}1.669365...) \\
&\leq \frac{1}{2}\log(e^{0.7402}) + \frac{1}{2}\log(\frac{a^{2(s+1)}\sigma_{v1}^2}{\sigma_{v2}^2}1.6694). \qquad (9.6)
\end{aligned}$$

(A): Assumptions $\widetilde{P_1} \leq \frac{\sigma_{v2}^2}{70a^{2(s-1)}}$ and $|a| \geq 2.5$.
(B): Assumption $\sigma_{v2}^2 \leq a^{2(s+1)}\sigma_{v1}^2$.

Moreover, since $\widetilde{P_1} \leq \frac{\sigma_{v2}^2}{70a^{2(s-1)}}$, we have

$$\frac{a^{2(k_2-k_1-1)}\widetilde{P_1}}{(1-2.5^{-1})^2} \leq \frac{a^{2s}}{(1-2.5^{-1})^2}\frac{\sigma_{v2}^2}{70a^{2(s-1)}} \leq \frac{a^2\sigma_{v2}^2}{(1-2.5^{-1})^2 70}. \qquad (9.7)$$

Therefore, by plugging (9.6) and (9.7) into (9.5), we have

$$\begin{aligned}
D_{L,1}(\widetilde{P_1},\widetilde{P_2}) &\geq (\sqrt{\frac{a^{2(s+1)}0.295a^2\sigma_{v1}^2}{e^{0.7402}\frac{a^{2(s+1)}\sigma_{v1}^2}{\sigma_{v2}^2}1.6694}} - \sqrt{\frac{a^2\sigma_{v2}^2}{(1-2.5^{-1})^2 70}})_+^2 + 1 \\
&\geq 0.008a^2\sigma_{v2}^2 + 1.
\end{aligned}$$

Proof of (b):

We choose $k_1, \Sigma, k_2, \alpha, \sigma_{v2}$ of Lemma 4.12 in the same way as the proof of (a) except k. Then, we will increase k arbitrarily large. Then, Lemma 4.12 reduces to

$$D_{L,1}(\widetilde{P_1},\widetilde{P_2}) \geq (\sqrt{\frac{a^{2(k-k_1)}\Sigma}{2^{2I'(\widetilde{P_1})}}} - \sqrt{\frac{a^{2(k-k_1-1)}\widetilde{P_1}}{(1-2.5^{-1})^2}} - \sqrt{\frac{a^{2(k-k_2-1)}2.5^{k_2-k_1}\widetilde{P_1}}{(1-2.5^{-1})^2}} - \sqrt{\frac{a^{2(k-k_2-1)}\widetilde{P_2}}{(1-2.5^{-1})^2}})_+^2 + 1. \qquad (9.8)$$

Since the relevant parameters are the same, (9.6) and (9.7) in the proof of (a) still hold.

Since $|a| \geq 2.5$ and $\widetilde{P_2} \leq \frac{a^4\sigma_{v2}^2}{28000}$, we also have

$$\begin{aligned}
&\frac{a^{2(k-k_2-1)}2.5^{k_2-k_1}\widetilde{P_1}}{(1-2.5^{-1})^2} \\
&= (\frac{2.5}{a^2})^{k_2-k_1}\frac{a^{2(k-k_1-1)}\widetilde{P_1}}{(1-2.5^{-1})^2} \\
&\leq (\frac{1}{2.5})^2\frac{a^{2(k-k_1-1)}\widetilde{P_1}}{(1-2.5^{-1})^2}
\end{aligned} \tag{9.9}$$

and

$$\frac{a^{2(k-k_2-1)}\widetilde{P_2}}{(1-2.5^{-1})^2} \leq \frac{a^{2(k-k_2-1)}}{(1-2.5^{-1})^2}\frac{a^4\sigma_{v2}^2}{28000}. \tag{9.10}$$

Therefore, by plugging (9.6), (9.7), (9.9), (9.10) into (9.8), we have

$$\begin{aligned}
D_{L,1}(\widetilde{P_1},\widetilde{P_2}) &\geq (\sqrt{\frac{a^{2(s+1)}0.295a^2\sigma_{v1}^2}{e^{0.7402\frac{a^{2(s+1)}\sigma_{v1}^2}{\sigma_{v2}^2}}1.6694}} - (1+\frac{1}{2.5})\sqrt{\frac{a^2\sigma_{v2}^2}{(1-2.5^{-1})^2 70}} - \sqrt{\frac{a^2\sigma_{v2}^2}{(1-2.5^{-1})^2 28000}})_+^2 a^{2(k-k_2)} + 1 \\
&\geq 10^{-6}a^{2(k-k_2+1)}\sigma_{v2}^2 + 1.
\end{aligned}$$

Since k can be chosen arbitrarily large, $\lim_{k\to\infty} D_{L,1}(\widetilde{P_1},\widetilde{P_2}) = \infty$.

Proof of (c):

We choose k, k_1, k_2, Σ of Lemma 4.12 in the same way as the proof of (a), i.e. $\Sigma = 0.295a^2\sigma_{v1}^2$ and $k_2 = k = k_1 + s + 1$. We put the remaining parameters α and σ'_{v2} as $\alpha = \frac{1}{2}$, $\sigma'^2_{v2} = 100a^{2(s-1)}\widetilde{P_1}$. Then, Lemma 4.12 reduces to

$$D_{L,1}(\widetilde{P_1},\widetilde{P_2}) \geq \frac{1}{2}D' + \frac{1}{2}D'' + 1$$

where

$$D' = (\sqrt{c\frac{a^{2(k_2-k_1)}\Sigma}{2^{2I'(\widetilde{P_1})}}} - \sqrt{c\frac{a^{2(k_2-k_1-1)}\widetilde{P_1}}{(1-2.5^{-1})^2}})_+^2 \tag{9.11}$$

$$D'' = (\sqrt{\frac{a^{2(k_2-k_1-1)}\Sigma}{2^{2I''(\widetilde{P_1})}}} - \sqrt{\frac{a^{2(k_2-k_1-2)}2.5\widetilde{P_1}}{(1-2.5^{-1})^2}})_+^2. \tag{9.12}$$

Here, $I''(\widetilde{P_1})$ is upper bounded as:

$$I''(\widetilde{P_1}) = \frac{1}{2}\log(1+\frac{1}{(k_2-k_1-1)\sigma_{v2}^2}(\frac{2a^{2(k_2-2-k_1)}}{1-a^{-2}}\Sigma+\frac{2a^{2(k_2-3-k_1)}}{1-a^{-2}}\frac{2.5\widetilde{P_1}}{(1-2.5a^{-2})(1-2.5^{-1})}))^{k_2-k_1-1}$$

$$\begin{aligned}
&\overset{(A)}{\leq} \frac{1}{2}\log(1+\frac{1}{(k_2-k_1-1)\sigma_{v2}^2}(\frac{2a^{2(k_2-2-k_1)}}{1-2.5^{-2}}0.295a^2\sigma_{v1}^2+\frac{2a^{2(k_2-3-k_1)}}{1-2.5^{-2}}\frac{2.5\frac{1}{20000}a^4\sigma_{v1}^2}{(1-2.5^{-1})^2}))^{k_2-k_1-1}\\
&= \frac{1}{2}\log(1+\frac{1}{(k_2-k_1-1)\sigma_{v2}^2}(\frac{2}{1-2.5^{-2}}0.295+\frac{2}{1-2.5^{-2}}\frac{2.5\frac{1}{20000}}{(1-2.5^{-1})^2})a^{2s}\sigma_{v1}^2)^{k_2-k_1-1}\\
&\overset{(B)}{\leq} \frac{1}{2}\log(1+\frac{1}{k_2-k_1-1}(\frac{2}{1-2.5^{-2}}0.295+\frac{2}{1-2.5^{-2}}\frac{2.5\frac{1}{20000}}{(1-2.5^{-1})^2}))^{k_2-k_1-1}\\
&= \frac{1}{2}\log(1+\frac{1}{k_2-k_1-1}0.703207...)^{k_2-k_1-1}\\
&\leq \frac{1}{2}\log(1+\frac{1}{k_2-k_1-1}0.7033)^{k_2-k_1-1}\\
&\leq \frac{1}{2}\log(e^{0.7033}).
\end{aligned}\tag{9.13}$$

(A): $\widetilde{P_1} \leq \frac{1}{20000}a^4\sigma_{v1}^2$ and $|a| \geq 2.5$.
(B): $a^{2s}\sigma_{v1}^2 \leq \sigma_{v2}^2$.

Likewise, $I'(\widetilde{P_1})$ is upper bounded as:

$$I'(\widetilde{P_1}) \leq \frac{1}{2}\log(1+\frac{1}{\sigma'^2_{v2}}(2a^{2(k_2-1-k_1)}\Sigma+2\frac{a^{2(k_2-2-k_1)}\widetilde{P_1}}{(1-2.5a^{-2})(1-2.5^{-1})}))+\frac{1}{2}\log(e^{0.7033})+\frac{1}{2}\log(\frac{2\pi e}{4})$$

$$\begin{aligned}
&\leq \frac{1}{2}\log(1+\frac{1}{\sigma'^2_{v2}}(2a^{2s}\Sigma+2\frac{a^{2(s-1)}\widetilde{P_1}}{(1-2.5^{-1})^2}))+\frac{1}{2}\log(e^{0.7033})+\frac{1}{2}\log(\frac{2\pi e}{4})\\
&= \frac{1}{2}\log(1+\frac{1}{\sigma'^2_{v2}}(2a^{2s}(0.295a^2\sigma_{v1}^2)+2\frac{a^{2(s-1)}\widetilde{P_1}}{(1-2.5^{-1})^2}))+\frac{1}{2}\log(e^{0.7033})+\frac{1}{2}\log(\frac{2\pi e}{4})\\
&= \frac{1}{2}\log(\frac{1}{100a^{2(s-1)}\widetilde{P_1}}(100a^{2(s-1)}\widetilde{P_1}+2\cdot 0.295a^{2(s+1)}\sigma_{v1}^2+2\frac{a^{2(s-1)}\widetilde{P_1}}{(1-2.5^{-1})^2}))+\frac{1}{2}\log(e^{0.7033})+\frac{1}{2}\log(\frac{2\pi e}{4})\\
&\overset{(A)}{\leq} \frac{1}{2}\log(\frac{1}{100a^{2(s-1)}\widetilde{P_1}}(\frac{100a^{2(s+1)}\sigma_{v1}^2}{20000}+2\cdot 0.295a^{2(s+1)}\sigma_{v1}^2+2\frac{a^{2(s+1)}\sigma_{v1}^2}{20000(1-2.5^{-1})^2}))\\
&+\frac{1}{2}\log(e^{0.7033})+\frac{1}{2}\log(\frac{2\pi e}{4})\\
&= \frac{1}{2}\log(\frac{1}{100}(0.595277...)\frac{a^{2(s+1)}\sigma_{v1}^2}{a^{2(s-1)}\widetilde{P_1}})+\frac{1}{2}\log(e^{0.7033})+\frac{1}{2}\log(\frac{2\pi e}{4})\\
&\leq \frac{1}{2}\log(\frac{1}{100}(0.5953)\frac{a^{2(s+1)}\sigma_{v1}^2}{a^{2(s-1)}\widetilde{P_1}})+\frac{1}{2}\log(e^{0.7033})+\frac{1}{2}\log(\frac{2\pi e}{4}).
\end{aligned}\tag{9.14}$$

(A): Assumption $\widetilde{P_1} \leq \frac{a^4\sigma_{v1}^2}{20000}$.

Therefore, by plugging (9.14) into (9.11), we get the following lower bound on D':

$$\begin{aligned}
D' &\geq c\Big(\sqrt{\frac{0.295a^{2(s+2)}\sigma_{v1}^2}{e^{0.7033}\frac{2\pi e}{4}\frac{1}{100}0.5953\frac{a^{2(s+1)}\sigma_{v1}^2}{a^{2(s-1)}\widetilde{P_1}}}} - \sqrt{\frac{a^{2s}\widetilde{P_1}}{(1-2.5^{-1})^2}}\Big)_+^2 \\
&= c\Big(\sqrt{\frac{0.295}{e^{0.7033}\frac{2\pi e}{4}\frac{1}{100}0.5953}} - \sqrt{\frac{1}{(1-2.5^{-1})^2}}\Big)_+^2 a^{2s}\widetilde{P_1} \\
&= 0.532969...ca^{2s}\widetilde{P_1} \\
&\geq 0.5329ca^{2s}\widetilde{P_1} \\
&= 0.5329\frac{2\cdot 10a^{s-1}\sqrt{\widetilde{P_1}}}{\sqrt{2\pi}\sigma_{v2}}\exp\Big(-\frac{100a^{2(s-1)}\widetilde{P_1}}{2\sigma_{v2}^2}\Big)a^{2s}\widetilde{P_1} \\
&\overset{(A)}{\geq} 0.5329\frac{2\cdot 10}{\sqrt{2\pi}\sqrt{70}}\exp\Big(-\frac{100a^{2(s-1)}\widetilde{P_1}}{2\sigma_{v2}^2}\Big)a^{2s}\widetilde{P_1} \\
&= 0.508202...a^{2s}\widetilde{P_1}\exp\Big(-\frac{50a^{2(s-1)}\widetilde{P_1}}{\sigma_{v2}^2}\Big) \\
&\geq 0.5082a^{2s}\widetilde{P_1}\exp\Big(-\frac{50a^{2(s-1)}\widetilde{P_1}}{\sigma_{v2}^2}\Big).
\end{aligned}$$

(A): Assumption $\widetilde{P_1} \geq \frac{\sigma_{v2}^2}{70a^{2(s-1)}}$.

Since $\widetilde{P_1} \leq \frac{a^4\sigma_{v1}^2}{20000}$, we also have

$$\begin{aligned}
&\frac{a^{2(k_2-k_1-2)}2.5\widetilde{P_1}}{(1-2.5^{-1})^2} \\
&\leq \frac{a^{2(s-1)}2.5\frac{a^4\sigma_{v1}^2}{20000}}{(1-2.5^{-1})^2} \\
&= \frac{2.5a^{2(s+1)}\sigma_{v1}^2}{20000(1-2.5^{-1})^2}
\end{aligned} \tag{9.15}$$

Therefore, by plugging (9.13) and (9.15) into (9.12), D'' is lower bounded as:

$$\begin{aligned}
D'' &\geq \Big(\sqrt{\frac{0.295a^{2(s+1)}\sigma_{v1}^2}{e^{0.7033}}} - \sqrt{\frac{2.5a^{2(s+1)}\sigma_{v1}^2}{20000(1-2.5^{-1})^2}}\Big)_+^2 \\
&= 0.132117...a^{2(s+1)}\sigma_{v1}^2 \\
&\geq 0.1321a^{2(s+1)}\sigma_{v1}^2
\end{aligned}$$

Finally, $D_{L,1}(\widetilde{P_1}, \widetilde{P_2})$ is lower bounded as:

$$D_{L,1}(\widetilde{P_1}, \widetilde{P_2}) \geq 0.2541a^{2s}\widetilde{P_1}\exp\Big(-\frac{50a^{2(s-1)}\widetilde{P_1}}{\sigma_{v2}^2}\Big) + 0.066a^{2(s+1)}\sigma_{v1}^2 + 1$$

Proof of (d):

We choose $k_1, k_2, \Sigma, \alpha, \sigma'_{v2}$ of Lemma 4.12 in the same way as the proof of (c) except k. k will be chosen arbitrarily large. Lemma 4.12 reduces to

$$D_{L,1}(\widetilde{P_1}, \widetilde{P_2}) \geq \frac{1}{2}a^{2(k-k_2)}D' + \frac{1}{2}a^{2(k-k_2)}D'' + 1 \tag{9.16}$$

where

$$D' = \left(\sqrt{c\frac{a^{2(k_2-k_1)}\Sigma}{2^{2I'(\widetilde{P_1})}}} - \sqrt{c\frac{a^{2(k_2-k_1-1)}\widetilde{P_1}}{(1-2.5^{-1})^2}} - \sqrt{\frac{a^{-2}2.5^{k_2-k_1}\widetilde{P_1}}{(1-2.5^{-1})^2}} - \sqrt{\frac{a^{-2}\widetilde{P_2}}{(1-2.5^{-1})^2}}\right)_+^2$$

$$D'' = \left(\sqrt{\frac{a^{2(k_2-k_1-1)}\Sigma}{2^{2I''(\widetilde{P_1})}}} - \sqrt{\frac{a^{2(k_2-k_1-2)}2.5\widetilde{P_1}}{(1-2.5^{-1})^2}} - \sqrt{\frac{a^{-2}\widetilde{P_2}}{(1-2.5^{-1})^2}}\right)_+^2.$$

Denote $P' := \sqrt{\frac{a^{-2}2.5^{k_2-k_1}\widetilde{P_1}}{(1-2.5^{-1})^2}} + \sqrt{\frac{a^{-2}\widetilde{P_2}}{(1-2.5^{-1})^2}}$. Then, following the same steps of the proof of (c), we can lower bound D' and D'' as follows:

$$D' \geq \left(\sqrt{0.5082a^{2s}\widetilde{P_1}\exp(-\frac{50a^{2(s-1)}\widetilde{P_1}}{\sigma_{v2}^2})} - \sqrt{P'}\right)_+^2 \tag{9.17}$$

$$\begin{aligned} D'' &\geq \left(\sqrt{0.1321a^{2(s+1)}\sigma_{v1}^2} - \sqrt{\frac{a^{-2}\widetilde{P_2}}{(1-2.5^{-1})^2}}\right)_+^2 \\ &\geq \left(\sqrt{0.1321a^{2(s+1)}\sigma_{v1}^2} - \sqrt{P'}\right)_+^2 \end{aligned} \tag{9.18}$$

Here, we have

$$\begin{aligned} &\frac{a^{-2}2.5^{k_2-k_1}\widetilde{P_1}}{(1-2.5^{-1})^2} \\ &\leq \frac{a^{-2}|a|^{s+1}\widetilde{P_1}}{(1-2.5^{-1})^2} \ (\because |a| \geq 2.5) \\ &\leq \frac{a^{-2}a^{2s}\widetilde{P_1}}{(1-2.5^{-1})^2} \ (\because s \geq 1) \\ &\leq \frac{a^{2s}\max(1, a^2\sigma_{v1}^2)}{20000(1-2.5^{-1})^2} \ (\because \widetilde{P_1} \leq \frac{\max(a^2, a^4\sigma_{v1}^2)}{20000}) \end{aligned}$$

Thus,

$$\begin{aligned} P' &\overset{(A)}{\leq} \sqrt{2(\frac{a^{-2}2.5^{k_2-k_1}\widetilde{P_1}}{(1-2.5^{-1})^2} + \frac{a^{-2}\widetilde{P_2}}{(1-2.5^{-1})^2})} \\ &\overset{(B)}{\leq} \sqrt{\frac{a^{2s}\max(1, a^2\sigma_{v1}^2)}{10000(1-2.5^{-1})^2} + \frac{2}{(1-2.5^{-1})^2}(0.0457a^{2s}\widetilde{P_1}\exp(-\frac{50a^{2(s-1)}\widetilde{P_1}}{\sigma_{v2}^2}) + 0.0113a^{2s}\max(1, a^2\sigma_{v1}^2))} \\ &\leq \sqrt{0.253889a^{2s}\widetilde{P_1}\exp(-\frac{50a^{2(s-1)}\widetilde{P_1}}{\sigma_{v2}^2}) + 0.0625a^{2s}\max(1, a^2\sigma_{v1}^2)} \end{aligned} \tag{9.19}$$

(A): Cauchy-Schwarz inequaility

(B): Assumptions on $\widetilde{P_1}$ and $\widetilde{P_2}$

By comparing (9.19) with (9.17) and (9.18), we can conclude that either D' or D'' has to be greater than 0. Moreover, since we can choose k arbitrarily large in (9.16), $\lim_{k\to\infty} D_{L,1}(\widetilde{P_1},\widetilde{P_2}) = \infty$.

Proof of (e):
The same as (a) of Corollary 4.1.

9.4 Proof of Corollary 4.3

For simplicity, we will only prove the case when $a \geq 2.5$. The proof for $a \leq -2.5$ easily follows by replacing a with $|a|$.

We will evaluate Lemma 4.7 with the parameters $w_1 = \frac{a^s d}{6}$ and $d = \sqrt{\frac{320000P}{a^2}}$. Then, we can easily see that $(d, w_1) \in S_{U,1}$. Furthermore, $\frac{a^s d}{\sigma_{v2}} \geq 13$ since

$$\begin{aligned}\frac{a^s d}{\sigma_{v2}} &= \frac{a^{s-1}\sqrt{320000P}}{\sigma_{v2}} \\ &\geq \frac{a^{s-1}}{\sigma_{v2}}\sqrt{4\cdot 80000\frac{\sigma_{v2}^2}{70a^{2(s-1)}}} \ (\because P \geq \frac{\sigma_{v2}^2}{70a^{2(s-1)}}) \\ &= \sqrt{\frac{4\cdot 80000}{70}} \geq 13.\end{aligned}$$

Then, we will upper bound $D_{U,1}(d, w_1)$. First, let's bound the second term of $D_{U,1}(d, w_1)$ in (4.5). The second term is upper bounded as

$$\begin{aligned}&\sum_{1\leq i\leq\infty} 4a^2(ia^s d + \frac{a^{s-1}d\frac{a}{a-1} + w_1}{2})^2 Q(\frac{(2i-1)a^s d - (a^{s-1}d\frac{a}{a-1} + w_1)}{2\sigma_{v2}}) \\ &= 4a^2(a^s d + \frac{a^{s-1}d\frac{a}{a-1} + w_1}{2})^2 Q(\frac{a^s d - (a^{s-1}d\frac{a}{a-1} + w_1)}{2\sigma_{v2}}) \\ &+ 4a^2(2a^s d + \frac{a^{s-1}d\frac{a}{a-1} + w_1}{2})^2 Q(\frac{3a^s d - (a^{s-1}d\frac{a}{a-1} + w_1)}{2\sigma_{v2}}) + \cdots \\ &= 4a^2(a^s d + \frac{a^{s-1}d\frac{a}{a-1} + \frac{a^s d}{6}}{2})^2 Q(\frac{a^s d - (a^{s-1}d\frac{a}{a-1} + \frac{a^s d}{6})}{2\sigma_{v2}}) \\ &+ 4a^2(2a^s d + \frac{a^{s-1}d\frac{a}{a-1} + \frac{a^s d}{6}}{2})^2 Q(\frac{3a^s d - (a^{s-1}d\frac{a}{a-1} + \frac{a^s d}{6})}{2\sigma_{v2}}) + \cdots \\ &\leq 4a^2(a^s d + \frac{1}{2(2.5-1)}a^s d + \frac{1}{12}a^s d)^2 Q(\frac{1}{2\sigma_{v2}}(1 - \frac{1}{2.5-1} - \frac{1}{6})a^s d) \\ &+ 4a^2(2a^s d + \frac{1}{2(2.5-1)}a^s d + \frac{1}{12}a^s d)^2 Q(\frac{1}{2\sigma_{v2}}(3 - \frac{1}{2.5-1} - \frac{1}{6})a^s d) + \cdots \\ &= 4a^2(a^s d)^2(1 + \frac{5}{12})^2 Q(\frac{1}{2\sigma_{v2}}(1 - \frac{5}{6})a^s d) \\ &+ 4a^2(a^s d)^2(2 + \frac{5}{12})^2 Q(\frac{1}{2\sigma_{v2}}(3 - \frac{5}{6})a^s d) + \cdots \end{aligned} \tag{9.20}$$

Denote $k := \frac{a^s d}{\sigma_{v2}}$. Since we already know $k \geq 13$, for all $n \geq 1$ we have

$$\begin{aligned}
&\frac{(n+\frac{5}{12})^2 Q(\frac{1}{2\sigma_{v2}}(2n-1-\frac{5}{6})a^s d)}{(n+1+\frac{5}{12})^2 Q(\frac{1}{2\sigma_{v2}}(2n+1-\frac{5}{6})a^s d)} \\
&\geq \frac{(n+\frac{5}{12})^2 (\frac{1}{\frac{1}{2\sigma_{v2}}(2n-1-\frac{5}{6})a^s d} - \frac{1}{(\frac{1}{2\sigma_{v2}}(2n-1-\frac{5}{6})a^s d)^3}) \exp(-\frac{(\frac{1}{2\sigma_{v2}}(2n-1-\frac{5}{6})a^s d)^2}{2})}{(n+1+\frac{5}{12})^2 (\frac{1}{\frac{1}{2\sigma_{v2}}(2n+1-\frac{5}{6})a^s d}) \exp(-\frac{(\frac{1}{2\sigma_{v2}}(2n+1-\frac{5}{6})a^s d)^2}{2})} \quad (\because \textit{Lemma } 4.5) \\
&= \frac{(n+\frac{5}{12})^2 (\frac{1}{\frac{1}{2}(2n-1-\frac{5}{6})k} - \frac{1}{(\frac{1}{2}(2n-1-\frac{5}{6})k)^3}) \exp(-\frac{(\frac{1}{2}(2n-1-\frac{5}{6})k)^2}{2})}{(n+1+\frac{5}{12})^2 (\frac{1}{\frac{1}{2}(2n+1-\frac{5}{6})k}) \exp(-\frac{(\frac{1}{2}(2n+1-\frac{5}{6})k)^2}{2})} \\
&\geq (\frac{1+\frac{5}{12}}{2+\frac{5}{12}})^2 \frac{(\frac{1}{\frac{1}{2}(2n-1-\frac{5}{6})k} - \frac{1}{(\frac{1}{2}(2n-1-\frac{5}{6})k)^3})}{(\frac{1}{\frac{1}{2}(2n+1-\frac{5}{6})k})} \frac{\exp(-\frac{(\frac{1}{12}k)^2}{2})}{\exp(-\frac{(\frac{13}{12}k)^2}{2})} \quad (\because n \geq 1) \\
&\geq (\frac{1+\frac{5}{12}}{2+\frac{5}{12}})^2 \exp(\frac{7}{12}k^2)(\frac{\frac{1}{2}(2n+1-\frac{5}{6})}{\frac{1}{2}(2n-1-\frac{5}{6})} - \frac{\frac{1}{2}(2n+1-\frac{5}{6})}{(\frac{1}{2}(2n-1-\frac{5}{6}))^3 13^2}) \quad (\because k \geq 13) \\
&\overset{(A)}{\geq} (\frac{1+\frac{5}{12}}{2+\frac{5}{12}})^2 \exp(\frac{7}{12}k^2) 0.99 \\
&\geq 10^{42} \quad (\because k \geq 13)
\end{aligned}$$

(A): When $n = 1$, we can check the inequality by computation. When $n \geq 2$, we have

$$\begin{aligned}
&\frac{\frac{1}{2}(2n+1-\frac{5}{6})}{\frac{1}{2}(2n-1-\frac{5}{6})} - \frac{\frac{1}{2}(2n+1-\frac{5}{6})}{(\frac{1}{2}(2n-1-\frac{5}{6}))^3 13^2} \\
&\geq 1 - \frac{\frac{1}{2}(4+1-\frac{5}{6})}{(\frac{1}{2}(4-1-\frac{5}{6}))^3 13^2} \geq 0.99.
\end{aligned}$$

Thus, the terms in (9.20) decrease faster than a geometric sequence with ratio 10^{-42} and thus can be upper bounded as

$$(9.20) \leq 4a^2(a^s d)^2(1+\frac{5}{12})^2 Q(\frac{1}{2\sigma_{v2}}(1-\frac{5}{6})a^s d)\frac{1}{1-10^{-42}}. \tag{9.21}$$

The third term of $D_{U,1}(d, w_1)$ in (4.5) can also be bounded similarly. We have

$$\begin{aligned}
&\frac{(a^s d + \frac{a^s d}{2})^2 \frac{1}{2}}{(2a^s d + \frac{a^s d}{2})^2 Q(\frac{a^s d}{\sigma_{v2}})} \geq \frac{(1+\frac{1}{2})^2}{(2+\frac{1}{2})^2 2Q(13)} \quad (\because \frac{a^s d}{\sigma_{v2}} \geq 13) \\
&\geq 10^{37}
\end{aligned}$$

and for $n \geq 2$

$$
\begin{aligned}
&\frac{(n+\frac{1}{2})^2 Q(\frac{(n-1)ad}{\sigma_{v2}})}{((n+1)+\frac{1}{2})^2 Q(\frac{nad}{\sigma_{v2}})} \\
&\geq \frac{(n+\frac{1}{2})^2 (\frac{1}{\frac{(n-1)ad}{\sigma_{v2}}} - \frac{1}{(\frac{(n-1)ad}{\sigma_{v2}})^3}) \exp(-\frac{(\frac{(n-1)ad}{\sigma_{v2}})^2}{2})}{((n+1)+\frac{1}{2})^2 (\frac{1}{\frac{nad}{\sigma_{v2}}}) \exp(-\frac{(\frac{nad}{\sigma_{v2}})^2}{2})} \ (\because \textit{Lemma } 4.5) \\
&= \exp(\frac{2n-1}{2} k^2) \frac{(n+\frac{1}{2})^2 (\frac{1}{(n-1)k} - \frac{1}{(n-1)^3 k^3})}{((n+1)+\frac{1}{2})^2 (\frac{1}{nk})} \\
&\geq \exp(\frac{3}{2} 13^2) \frac{(2+\frac{1}{2})^2}{(3+\frac{1}{2})^2} (\frac{n}{n-1} - \frac{n}{(n-1)^3 13^2}) \ (\because n \geq 2, k \geq 13) \\
&\overset{(A)}{\geq} \exp(\frac{3}{2} 13^2) \frac{(2+\frac{1}{2})^2}{(3+\frac{1}{2})^2} 0.98 \\
&\geq 10^{109}.
\end{aligned}
$$

(A): Since $n \geq 2$, we have

$$
\frac{n}{n-1} - \frac{n}{(n-1)^3 13^2} \geq 1 - \frac{2}{(2-1)^3 13^2} \geq 0.98.
$$

Therefore, the third term of $D_{U,1}(d, w_1)$ in (4.5) is upper bounded by

$$
4a^2 Q(\frac{\frac{a^s d}{6}}{2\sqrt{a^{2(s-1)} \frac{a^2}{a^2-1} + a^{2s}\sigma_{v1}^2}})(a^s d + \frac{a^s d}{2})^2 \frac{1}{1-10^{-37}}. \tag{9.22}
$$

By plugging (9.21) and (9.22) into (4.5), we can bound $D_{U,1}(d, w_1)$ of Lemma 4.7 as follows.

$$
\begin{aligned}
D_{U,1}(d,w_1) &\le 2a^{2s}(2(\frac{d}{2})^2(\frac{1}{1-\frac{1}{a}})^2 + 2(\frac{1}{1-\frac{1}{a^2}}) + 2a^2\sigma_{v1}^2) \\
&+ 4a^2(a^s d)^2(1+\frac{5}{12})^2 Q(\frac{1}{2\sigma_{v2}}(1-\frac{5}{6})a^s d)\frac{1}{1-10^{-42}} \\
&+ 4a^2 Q(\frac{\frac{a^s d}{6}}{2\sqrt{a^{2(s-1)}\frac{a^2}{a^2-1} + a^{2s}\sigma_{v1}^2}})(a^s d + \frac{a^s d}{2})^2\frac{1}{1-10^{-37}} \\
&+ 2(a^2(\frac{d}{2})^2) + 1 \\
&\overset{(A)}{\le} 2a^{2s}(2(\frac{d}{2})^2(\frac{5}{3})^2 + \frac{50}{21} + 2a^2\sigma_{v1}^2) \\
&+ 4a^2(a^s d)^2(\frac{17}{12})^2 Q(\frac{a^s d}{12\sigma_{v2}})\frac{1}{1-10^{-42}} \\
&+ 4a^2 Q(\frac{\frac{a^s d}{6}}{2\sqrt{a^{2(s-1)}\frac{25}{21} + a^{2s}\sigma_{v1}^2}})(a^s d + \frac{a^s d}{2})^2\frac{1}{1-10^{-37}} \\
&+ 2(a^2(\frac{d}{2})^2) + 1 \\
&\overset{(B)}{\le} 2a^{2s}(2(\frac{d}{2})^2(\frac{5}{3})^2 + \frac{50}{21} + 2a^2\sigma_{v1}^2) \\
&+ 4a^2(a^s d)^2(\frac{17}{12})^2 Q(\frac{a^s d}{12\sigma_{v2}})\frac{1}{1-10^{-42}} \\
&+ 4a^2(a^s d + \frac{a^s d}{2})^2 Q(\frac{a^s d}{12\sqrt{\frac{46}{21}}\sigma_{v2}})\frac{1}{1-10^{-37}} \\
&+ 2(a^2(\frac{d}{2})^2) + 1 \\
&\le 2a^{2s}(2(\frac{d}{2})^2(\frac{5}{3})^2 + \frac{50}{21} + 2a^2\sigma_{v1}^2) \\
&+ 4a^2(a^s d)^2(\frac{17}{12})^2 Q(\frac{a^s d}{12\sqrt{\frac{46}{21}}\sigma_{v2}})\frac{1}{1-10^{-42}} \\
&+ 4a^2(a^s d + \frac{a^s d}{2})^2 Q(\frac{a^s d}{12\sqrt{\frac{46}{21}}\sigma_{v2}})\frac{1}{1-10^{-37}} \\
&+ 2(a^2(\frac{d}{2})^2) + 1 \\
&= 1 + \frac{100}{21}a^{2s} + \frac{25}{9}a^{2s}d^2 + \frac{a^2 d^2}{2} + 4a^{2(s+1)}\sigma_{v1}^2 \\
&+ (4(\frac{17}{12})^2\frac{1}{1-10^{-42}} + 9\frac{1}{1-10^{-37}})a^{2(s+1)}d^2 Q(\frac{a^s d}{12\sqrt{\frac{46}{21}}\sigma_{v2}}) \\
&\le 1 + \frac{100}{21}a^{2s} + \frac{25}{9}a^{2s}d^2 + \frac{a^2 d^2}{2} + 4a^{2(s+1)}\sigma_{v1}^2 + 17.03a^{2(s+1)}d^2 Q(\frac{a^s d}{12\sqrt{\frac{46}{21}}\sigma_{v2}})
\end{aligned}
$$

$$= 1 + \frac{100}{21}a^{2s} + \frac{25}{9}4 \cdot 80000a^{2(s-1)}P + 2 \cdot 80000P + 4a^{2(s+1)}\sigma_{v1}^2$$
$$+ 17.03 \cdot 4 \cdot 80000a^{2s}P \cdot Q(\frac{a^{s-1}\sqrt{4 \cdot 80000P}}{12\sqrt{\frac{46}{21}}\sigma_{v2}})$$
$$\overset{(C)}{\leq} 1 + \frac{100}{21}a^{2s} + \frac{25}{9}16a^{2(s-1)}\max(a^2, a^4\sigma_{v1}^2) + 8\max(a^2, a^4\sigma_{v1}^2) + 4a^{2(s+1)}\sigma_{v1}^2$$
$$+ 17.03 \cdot 4 \cdot 80000a^{2s}P\frac{1}{\sqrt{2\pi}\frac{a^{s-1}\sqrt{4 \cdot 80000P}}{12\sqrt{\frac{46}{21}}\sigma_{v2}}}\exp(-\frac{1}{2}\frac{a^{2(s-1)}4 \cdot 80000P}{144 \cdot \frac{46}{21}\sigma_{v2}^2})$$
$$= 1 + \frac{100}{21}a^{2s} + \frac{25}{9}16a^{2(s-1)}\max(a^2, a^4\sigma_{v1}^2) + 8\max(a^2, a^4\sigma_{v1}^2) + 4a^{2(s+1)}\sigma_{v1}^2$$
$$+ 17.03 \cdot 4 \cdot 80000\frac{\sqrt{70}}{\sqrt{2\pi}\frac{\sqrt{4 \cdot 80000}}{12\sqrt{\frac{46}{21}}}}\exp(-(\frac{1}{2}\frac{4 \cdot 80000}{144 \cdot \frac{46}{21}} - 50)\frac{a^{2(s-1)}P}{\sigma_{v2}^2})a^{2s}P\exp(-\frac{50a^{2(s-1)}P}{\sigma_{v2}^2})$$
$$\overset{(D)}{\leq} 1 + \frac{100}{21}a^{2s} + \frac{25}{9}16a^{2(s-1)}\max(a^2, a^4\sigma_{v1}^2) + 8\max(a^2, a^4\sigma_{v1}^2) + 4a^{2(s+1)}\sigma_{v1}^2$$
$$+ 17.03 \cdot 4 \cdot 80000\frac{\sqrt{70}}{\sqrt{2\pi}\frac{\sqrt{4 \cdot 80000}}{12\sqrt{\frac{46}{21}}}}\exp(-(\frac{1}{2}\frac{4 \cdot 80000}{144 \cdot \frac{46}{21}} - 50)\frac{1}{70})a^{2s}P\exp(-\frac{50a^{2(s-1)}P}{\sigma_{v2}^2})$$
$$\leq 1 + \frac{100}{21}a^{2s} + \frac{25}{9}16a^{2(s-1)}\max(a^2, a^4\sigma_{v1}^2) + 8\max(a^2, a^4\sigma_{v1}^2) + 4a^{2(s+1)}\sigma_{v1}^2$$
$$+ 831.473...a^{2s}P\exp(-\frac{50a^{2(s-1)}P}{\sigma_{v2}^2})$$
$$\leq 1 + (\frac{100}{21} + \frac{25}{9}16 + 8 + 4)a^{2s}\max(1, a^2\sigma_{v1}^2) + 831.473...a^{2s}P\exp(-\frac{50a^{2(s-1)}P}{\sigma_{v2}^2})$$
$$\leq 1 + 61.206...a^{2s}\max(1, a^2\sigma_{v1}^2) + 831.473...a^{2s}P\exp(-\frac{50a^{2(s-1)}P}{\sigma_{v2}^2})$$
$$\leq 1 + 62a^{2s}\max(1, a^2\sigma_{v1}^2) + 832a^{2s}P\exp(-\frac{50a^{2(s-1)}P}{\sigma_{v2}^2})$$
$$\leq 832a^{2s}P\exp(-\frac{50a^{2(s-1)}P}{\sigma_{v2}^2}) + 63a^{2s}\max(1, a^2\sigma_{v1}^2)$$

(A): $a \geq 2.5$.

(B): From the assumption $a^{2(s-1)}\max(1, a^2\sigma_{v1}^2) \leq \sigma_{v2}^2 \leq a^{2s}\max(1, a^2\sigma_{v1}^2)$, we have

$$a^{2(s-1)}\frac{25}{21} + a^{2s}\sigma_{v1}^2 \leq \frac{46}{21}\max(a^{2(s-1)}, a^{2s}\sigma_{v1}^2) \leq \frac{46}{21}\sigma_{v2}^2.$$

(C): $P \leq \frac{\max(a^2, a^4\sigma_{v1}^2)}{20000}$ and Lemma 4.5.

(D): $P \geq \frac{\sigma_{v2}^2}{70a^{2(s-1)}}$.

This justifies the upper bound on $D_U(P_1, P_2)$. By the definition of d and Lemma 4.7, P_1

is upper bounded by $80000P$. P_2 of Lemma 4.7 can be upper bounded as

$$\begin{aligned}
P_2 &\leq 8a^2 D_{U,1}(d, w_1) + \frac{7}{2}a^{2(s+1)}d^2 + 4a^2\sigma_{v2}^2 \\
&= 8a^2 D_{U,1}(d, w_1) + \frac{7}{2}4 \cdot 80000a^{2s}P + 4a^2\sigma_{v2}^2 \\
&\leq 8a^2(832a^{2s}P\exp(-\frac{50a^{2(s-1)}P}{\sigma_{v2}^2}) + 63a^{2s}\max(1, a^2\sigma_{v1}^2)) \\
&+ \frac{7}{2}4 \cdot 4a^{2s}\max(a^2, a^4\sigma_{v1}^2) + 4a^2a^{2s}\max(1, a^2\sigma_{v1}^2) \\
&= 8a^2(832a^{2s}P\exp(-\frac{50a^{2(s-1)}P}{\sigma_{v2}^2}) + 70.5a^{2s}\max(1, a^2\sigma_{v1}^2)) \\
&= 6656a^{2(s+1)}P\exp(-\frac{50a^{2(s-1)}P}{\sigma_{v2}^2}) + 564a^{2(s+1)}\max(1, a^2\sigma_{v1}^2).
\end{aligned}$$

where the inequality comes from the assumptions $P \leq \frac{\max(a^2, a^4\sigma_{v1}^2)}{20000}$ and $\sigma_{v2}^2 \leq a^{2s}\max(1, a^2\sigma_{v1}^2)$. This finishes the proof.

9.5 Proof of Proposition 4.8

As the proof of Proposition 4.7, by Lemma 4.14 it is enough to show that there exists $c \geq 1$ such that $D_U(c\widetilde{P_1}, c\widetilde{P_2}) \leq c \cdot D_L(\widetilde{P_1}, \widetilde{P_2})$.

(i) When $\widetilde{P_1} \leq \frac{\sigma_{v2}^2}{70a^{2(s-1)}}$ and $\widetilde{P_2} \leq \frac{a^4\sigma_{v2}^2}{28000}$

Lower bound: By Corollary 4.2 (b)

$$D_L(\widetilde{P_1}, \widetilde{P_2}) = \infty.$$

Therefore, we do not need the corresponding upper bound.

(ii) When $\widetilde{P_1} \leq \frac{\sigma_{v2}^2}{70a^{2(s-1)}}$ and $\widetilde{P_2} \geq \frac{a^4\sigma_{v2}^2}{28000}$

Lower bound: By Corollary 4.2 (a)

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq 0.008a^2\sigma_{v2}^2 + 1.$$

Upper bound: By Lemma 4.15

$$\begin{aligned}
(D_U(P_1, P_2), P_1, P_2) &\leq (a^2\sigma_{v2}^2 + 1, 0, a^4\sigma_{v2}^2 + a^2\sigma_{v2}^2 + a^2) \\
&\leq (a^2\sigma_{v2}^2 + 1, 0, a^4\sigma_{v2}^2 + a^2\sigma_{v2}^2 + a^2\sigma_{v2}^2) \\
&\leq (a^2\sigma_{v2}^2 + 1, 0, (1 + \frac{2}{2.5^2})a^4\sigma_{v2}^2) \\
&\leq (a^2\sigma_{v2}^2 + 1, 0, 1.32a^4\sigma_{v2}^2).
\end{aligned}$$

Ratio: Thus, c is upper bounded by

$$c \leq \max(\frac{1}{0.008}, \frac{1.32}{\frac{1}{28000}}).$$

(iii) When $\frac{\sigma_{v2}^2}{70a^{2(s-1)}} \leq \widetilde{P_1} \leq \frac{1}{20000}\max(a^2, a^4\sigma_{v1}^2)$ and $\widetilde{P_2} \leq 0.0457a^{2(s+1)}\widetilde{P_1}\exp(-\frac{50a^{2(s-1)}\widetilde{P_1}}{\sigma_{v2}^2})+$ $0.0113a^{2(s+1)}\max(1, a^2\sigma_{v1}^2)$

Lower bound: By Corollary 4.2 (d)

$$D_L(\widetilde{P_1}, \widetilde{P_2}) = \infty$$

Therefore, we do not need the corresponding upper bound.

(iv) When $\frac{\sigma_{v2}^2}{70a^{2(s-1)}} \leq \widetilde{P_1} \leq \frac{1}{20000}\max(a^2, a^4\sigma_{v1}^2)$ and $\widetilde{P_2} \geq 0.0457a^{2(s+1)}\widetilde{P_1}\exp(-\frac{50a^{2(s-1)}\widetilde{P_1}}{\sigma_{v2}^2})+$ $0.0113a^{2(s+1)}\max(1, a^2\sigma_{v1}^2)$

Lower bound: By Corollary 4.2 (c)

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq 0.2541a^{2s}\widetilde{P_1}\exp(-\frac{50a^{2(s-1)}\widetilde{P_1}}{\sigma_{v2}^2}) + 0.066a^{2s}\max(1, a^2\sigma_{v1}^2) + 1$$

Upper bound: By Corollary 4.3

$$\begin{aligned}(D_U(P_1, P_2), P_1, P_2) \leq &(63a^{2s}\max(1, a^2\sigma_{v1}^2) + 832a^{2s}\widetilde{P_1}\exp(-\frac{50a^{2(s-1)}\widetilde{P_1}}{\sigma_{v2}^2}), 80000\widetilde{P_1} \\ &, 6656a^{2(s+1)}\widetilde{P_1}\exp(-\frac{50a^{2(s-1)}\widetilde{P_1}}{\sigma_{v2}^2}) + 564a^{2(s+1)}\max(1, a^2\sigma_{v1}^2))\end{aligned}$$

Ratio: c is upper bounded by

$$c \leq \max(\frac{832}{0.2541}, \frac{63}{0.066}, 80000, \frac{6656}{0.0457}, \frac{564}{0.0113})$$

(v) When $\widetilde{P_1} \geq \frac{1}{20000}\max(a^2, a^4\sigma_{v1}^2)$

Lower bound: By Corollary 4.2 (e)

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq 0.295 \cdot \max(1, a^2\sigma_{v1}^2)$$

Upper bound: By Lemma 4.15

$$\begin{aligned}(D_U(P_1, P_2), P_1, P_2) &\leq (a^2\sigma_{v1}^2 + 1, a^4\sigma_{v1}^2 + a^2\sigma_{v1}^2 + a^2, 0) \\ &\leq (a^2\sigma_{v1}^2 + 1, 2a^4\sigma_{v1}^2 + a^2, 0) \\ &\leq (2\max(1, a^2\sigma_{v1}^2), 3\max(a^2, a^4\sigma_{v1}^2), 0)\end{aligned}$$

Ratio: c is upper bounded by

$$c \leq \max(\frac{2}{0.295}, \frac{3}{\frac{1}{20000}})$$

Therefore, by (i), (ii), (iii), (iv), (v), the lemma is true and $c \leq 1.5 \times 10^5$.

9.6 Proof of Proposition 4.4

Since a goes to infinity, let $a \geq 10000$. We will first show the best linear strategy performance is $\Theta(a^3)$.

• The best linear strategy performance is $\Theta(a^3)$: Following the same steps in the proof of Lemma 4.8, we can lower bound the average cost as follows:

$$
\begin{aligned}
&\inf_{u_1,u_2 \in L'_{lin}} \frac{1}{N} \sum_{0 \le n < N} \mathbb{E}[qx^2[n] + r_1 u_1^2[n]] \\
&= \inf_{u_1,u_2 \in L'_{lin}} \frac{1}{N}((\frac{1}{2}r_1\mathbb{E}[u_1^2[1]] + \frac{1}{2}r_1\mathbb{E}[u_1^2[2]] + q\mathbb{E}[x^2[3]]) + (\frac{1}{2}r_1\mathbb{E}[u_1^2[2]] + \frac{1}{2}r_1\mathbb{E}[u_1^2[3]] + q\mathbb{E}[x^2[4]]) + \cdots \\
&+ (\frac{1}{2}r_1\mathbb{E}[u_1^2[N-3]] + \frac{1}{2}r_1\mathbb{E}[u_1^2[N-2]] + q\mathbb{E}[x^2[N-1]])) \\
&\ge \frac{N-3}{N} \inf_{u_1,u_2 \in L'_{lin}} (\frac{1}{2}r_1\mathbb{E}[u_1^2[1]] + \frac{1}{2}r_1\mathbb{E}[u_1^2[2]] + q\mathbb{E}[x^2[3]]) \\
&= \frac{N-3}{N} \inf_{u_1,u_2 \in L'_{lin}} (\frac{1}{2}a\mathbb{E}[u_1^2[1]] + \frac{1}{2}a\mathbb{E}[u_1^2[2]] + \mathbb{E}[x^2[3]]).
\end{aligned}
$$

In the similar way of Proposition 4.6, we can further justify that setting $w[1] = 0$, $w[2] = 0$ only decrease the quadratic cost. Then, at time 1 we have

$$
\begin{aligned}
x[1] &= w[0] \\
y_1[1] &= w[0] \\
y_2[1] &= w[0] + v_2[1]
\end{aligned}
$$

Let

$$
\begin{aligned}
u_1[1] &= k_{11}w[0] \\
u_2[1] &= k_{21}(w[0] + v_2[1])
\end{aligned}
$$

At time 2 we have

$$
\begin{aligned}
x[2] &= ax[1] + u_1[1] + u_2[1] \\
&= aw[0] + k_{11}w[0] + k_{21}(w[0] + v_2[1]) \\
y_1[1] &= aw[0] + k_{11}w[0] + k_{21}(w[0] + v_2[1]) \\
y_2[1] &= aw[0] + k_{11}w[0] + k_{21}(w[0] + v_2[1]) + v_2[2]
\end{aligned}
$$

Therefore, we can put

$$
\begin{aligned}
u_1[2] &= k_{12}w[0] + k_{13}v_2[1] \\
u_2[3] &= k_{22}(w[0] + v_2[1]) + k_{23}(aw[0] + k_{11}w[0] + v_2[2])
\end{aligned}
$$

At time 3 we have

$$\begin{aligned}
x[3] &= ax[2] + u_1[2] + u_2[2] \\
&= a^2 w[0] + ak_{11}w[0] + ak_{21}(w[0] + v_2[1]) + k_{12}w[0] + k_{13}v_2[1] \\
&+ k_{22}(w[0] + v_2[1]) + k_{23}(aw[0] + k_{11}w[0] + v_2[2]) \\
&= (a^2 + ak_{11} + k_{12})w[0] + k_{13}v_2[1] + (ak_{21} + k_{22})(w[0] + v_2[1]) + k_{23}(aw[0] + k_{11}w[0] + v_2[2]) \\
&= (a^2 + ak_{11} + k_{12} - k_{13})w[0] + (ak_{21} + k_{22} + k_{13})(w[0] + v_2[1]) + k_{23}(aw[0] + k_{11}w[0] + v_2[2])
\end{aligned}$$

Therefore,

$$\mathbb{E}[x^2[3]] \geq (a^2 + ak_{11} + k_{12} - k_{13})^2 MMSE[w[0]|w[0] + v_2[1], aw[0] + k_{11}w[0] + v_2[2]]$$

(i) When $\mathbb{E}[u_1^2[1]] + \mathbb{E}[u_1^2[2]] \leq \frac{1}{16}a^2$

The condition implies

$$\begin{aligned}
&\mathbb{E}[(k_{11}w[0])^2] + \mathbb{E}[(k_{12}w[0] + k_{13}v_2[1])^2] \\
&= k_{11}^2 + k_{12}^2 + k_{13}^2 a \leq \frac{1}{16}a^2
\end{aligned}$$

Thus,

$$\begin{aligned}
|k_{11}| &\leq \frac{1}{4}a \\
|k_{12}| &\leq \frac{1}{4}a \\
|k_{13}| &\leq \frac{1}{4}\sqrt{a}
\end{aligned}$$

Since $a \geq 10000$ we have

$$a^2 + ak_{11} + k_{12} - k_{13} \geq a^2 - \frac{a^2}{4} - \frac{a^2}{4} - \frac{a^2}{4} = \frac{a^2}{4}$$

Moreover, we also have

$$\begin{aligned}
&MMSE[w[0]|w[0] + v_2[1], aw[0] + k_{11}w[0] + v_2[2]] \\
&\geq MMSE[w[0]|w[0] + v_2[1], \frac{5a}{4}w[0] + v_2[2]] \\
&\geq MMSE[w[0]|\frac{5a}{4}w[0] + v_2[1], \frac{5a}{4}w[0] + v_2[2]] \\
&= MMSE[w[0]|\frac{10a}{4}w[0] + v_2[1] + v_2[2]] \\
&= 1 - \frac{(\frac{10a}{4})^2}{(\frac{10a}{4})^2 + 2a} \\
&= \frac{2a}{(\frac{10a}{4})^2 + 2a} \\
&\geq \frac{2}{(\frac{10}{4})^2 + 2}\frac{1}{a} = \frac{8}{33a}
\end{aligned}$$

Therefore, in this case,

$$\inf_{u_1,u_2 \in L'_{lin}} \frac{1}{2}a\mathbb{E}[u_1^2[1]] + \frac{1}{2}a\mathbb{E}[u_1^2[2]] + \mathbb{E}[x^2[3]] \geq \frac{1}{16} \cdot \frac{8}{33}a^3$$

(ii) When $\mathbb{E}[u_1^2[1]] + \mathbb{E}[u_1^2[2]] \geq \frac{1}{16}a^2$ In this case

$$\inf_{u_1,u_2 \in L'_{lin}} \frac{1}{2}a\mathbb{E}[u_1^2[1]] + \frac{1}{2}a\mathbb{E}[u_1^2[2]] + \mathbb{E}[x^2[3]] \geq \frac{1}{32}a^3$$

Therefore, by (i),(ii),

$$\inf_{u_1,u_2 \in L'_{lin}} \frac{1}{2}a\mathbb{E}[u_1^2[1]] + \frac{1}{2}a\mathbb{E}[u_1^2[2]] + \mathbb{E}[x^2[3]] \geq \frac{1}{66}a^3 \tag{9.23}$$

• The optimal average cost is $O(a^2 \log a)$: Now, we will show the average cost, $O(a^2 \log a)$, is achievable by the nonlinear 1-stage signaling strategy. Let $a \geq 20000$. Since $\max(1, a^2\sigma_{v1}^2) \leq \sigma_{v2}^2 \leq a^2 \max(1, a^2\sigma_{v1}^2)$ and $\frac{a}{70} \leq \frac{a \log a}{25} \leq \frac{1}{20000}a^2$, we can set $s = 1$ and $P = \frac{a \log a}{25}$ in Corollary 4.2. Then, by Corollary 4.2, the average cost is upper bounded as follows.

$$\begin{aligned}
&\inf_{u_1,u_2 \in L_{sig,1}} \frac{1}{N} \sum_{0 \leq n < N} \mathbb{E}[qx^2[n] + r_1 u_1^2[n]] \\
&\leq 832a^2 \frac{a \log a}{25} \exp(-\frac{50\frac{a \log a}{25}}{a}) + 63a^2 + a\frac{80000a \log a}{25} \\
&\leq 832\frac{a \log a}{25} + 63a^2 + \frac{80000}{25}a^2 \log a \\
&\leq 3297a^2 \log a
\end{aligned} \tag{9.24}$$

In short, by (9.23) the optimal linear strategy cost is lower bounded by $\Omega(a^3)$. By (9.24), the nonlinear 1-stage signaling strategy can achieve $O(a^2 \log a)$. Thus, their ratio diverges as a goes to infinity, which finishes the proof.

Chapter 10

Appendix for Chapter 5

10.1 Proof of Corollary 5.1, 5.2, 5.3

Proof of Corollary 5.1 of Page 230. For simplicity, we will only proof for the case when $a = 1$. The proof for the case of $a = -1$ follows similarly by replacing a with $-a$.

In this case, Lemma 5.1 reduces to that for all $|1 - k| < 1$,

$$D_{\sigma_v}(P) \leq \frac{(2k - k^2)\Sigma_E + 1}{1 - (1 - k)^2} = \frac{(2k - k^2)\Sigma_E + 1}{2k - k^2} = \frac{1}{2k - k^2} + \Sigma_E \tag{10.1}$$

$$P \leq k^2(\frac{(2k - k^2)\Sigma_E + 1}{1 - (1 - k)^2} - \Sigma_E) = k^2(\frac{1}{2k - k^2} + \Sigma_E - \Sigma_E) = \frac{k^2}{2k - k^2} \tag{10.2}$$

where

$$\Sigma_E = \frac{-1 + \sqrt{4\sigma_v^2 + 1}}{2}.$$

Let $k^\star \in (0, 1]$ be a constant such that $\max(1, \Sigma_E) = \frac{1}{2k^\star - k^{\star 2}}$. Here, we can see that such $k^\star$ always exists since $\max(1, \Sigma_E) \geq 1$ and $\frac{1}{2k-k^2}$ is a decreasing function on k. Let $k \in (0, k^\star]$. Then, we can see since $0 < k^\star \leq 1$, $|1 - k| < 1$. Then, (10.1) and (10.2) are again upper bounded as follows:

$$\begin{aligned} D_{\sigma_v}(P) &= \frac{1}{2k - k^2} + \Sigma_E \\ &\leq \frac{1}{2k - k^2} + \max(1, \Sigma_E) \\ &= \frac{1}{2k - k^2} + \frac{1}{2k^\star - k^{\star 2}} \\ &\leq \frac{2}{2k - k^2} \end{aligned}$$

where the last inequality follows from $0 < k \leq k^\star$.

$$\begin{aligned} P &= \frac{k^2}{2k - k^2} \\ &\leq \frac{k^2(2-k)^2}{2k-k^2} \\ &= 2k - k^2 \end{aligned}$$

where the inequality follows from $0 < k \leq k^\star \leq 1$.

Let's put $t = 2k - k^2$. Then, we have $(D_{\sigma_v}(P), P) \leq (\frac{2}{t}, t)$ where $t \in (0, 2k^\star - k^{\star 2}]$. Therefore, $t \in (0, \frac{1}{\max(1,\Sigma_E)}]$. This finishes the proof of the first claim.

When $\sigma_v \geq 16$, we have

$$\begin{aligned} \Sigma_E &= \frac{-1 + \sqrt{4\sigma_v^2 + 1}}{2} \leq \frac{\sqrt{4\sigma_v^2+1}}{2} \\ &\leq \frac{\sqrt{4\sigma_v^2 + \frac{1}{16^2}\sigma_v^2}}{2} = 1.000488...\sigma_v \\ &\leq 1.0005\sigma_v. \end{aligned}$$

Therefore, the range of t at least includes $(0, \frac{1}{1.0005\sigma_v}]$ and the second claim is true.

When $\sigma_v \leq 16$, we have

$$\Sigma_E = \frac{-1 + \sqrt{4 \cdot 16^2 + 1}}{2} = 15.0078105... \leq 15.008.$$

Therefore,the range of t at least includes $(0, \frac{1}{15.008}]$ and the third claim is true. □

Proof of Corollary 5.2 of Page 233. For simplicity, we prove only for the case when $a > 1$. The proof for the case of $a < -1$ follows similarly by replacing a with $-a$.

Proof of (i): Let's put $k = a - \frac{1}{a}$ in Lemma 5.1. Since $|a - a + \frac{1}{a}| = |\frac{1}{a}| < 1$, the power-distortion tradeoff in (5.3) still holds. Thus, we can see that

$$\begin{aligned} D_U(P) &\leq \frac{(2a(a-\frac{1}{a}) - (a - \frac{1}{a})^2)\Sigma_E + 1}{1 - (\frac{1}{a})^2} \\ &= \frac{a^2 - \frac{1}{a^2}}{1 - \frac{1}{a^2}}\Sigma_E + \frac{1}{1-(\frac{1}{a})^2} \\ &= (a^2+1)\Sigma_E + \frac{a^2}{a^2-1} \end{aligned}$$

and

$$\begin{aligned} P &\leq (\frac{a^2-1}{a})^2(a^2\Sigma_E + \frac{a^2}{a^2-1}) - \Sigma_E \\ &\leq (a^2-1)^2\Sigma_E + (a^2-1), \end{aligned}$$

which finishes the proof of (i).

Proof of (ii): We will divide into two cases depending on Σ_E.

Case 1) When $\max(1, (1+a^2)\Sigma_E) > \frac{1}{1-(\frac{1}{a})^2}$.
In this case, the domain for t is an empty set and we do not have to prove anything.
Case 2) When $\max(1, (1+a^2)\Sigma_E) \leq \frac{1}{1-(\frac{1}{a})^2}$.
Since $\max(1, (1+a^2)\Sigma_E) \leq \frac{1}{1-(\frac{1}{a})^2}$, there exists $\Delta^\star \in [0, \frac{1}{a}]$ such that

$$\max(1, (a^2+1)\Sigma_E) = \frac{1}{1-(\frac{1}{a}-\Delta^\star)^2}.$$

Let's put $k = a - \frac{1}{a} + \Delta$ in Lemma 5.1 where $\Delta \in [0, \Delta^\star]$. Then, we have the following upper bound on $D_{\sigma_v}(P)$ and P.

$$\begin{aligned}
D_{\sigma_v}(P) &\leq \frac{(2ak-k^2)\Sigma_E + 1}{1-(a-k)^2} && (10.3)\\
&= \frac{2ak-k^2}{1-(a-k)^2}\Sigma_E + \frac{1}{1-(a-k)^2}\\
&= \frac{a^2-1+1-(a-k)^2}{1-(a-k)^2}\Sigma_E + \frac{1}{1-(a-k)^2}\\
&= (\frac{a^2-1}{1-(a-k)^2}+1)\Sigma_E + \frac{1}{1-(a-k)^2}\\
&= (\frac{a^2-1}{1-(\frac{1}{a}-\Delta)^2}+1)\Sigma_E + \frac{1}{1-(a-k)^2}\\
&\overset{(A)}{\leq} (\frac{a^2-1}{1-(\frac{1}{a})^2}+1)\Sigma_E + \frac{1}{1-(a-k)^2}\\
&= (a^2+1)\Sigma_E + \frac{1}{1-(a-k)^2}\\
&= (a^2+1)\Sigma_E + \frac{1}{1-(\frac{1}{a}-\Delta)^2}\\
&\overset{(B)}{\leq} \max(1, (a^2+1)\Sigma_E) + \frac{1}{1-(\frac{1}{a}-\Delta)^2}\\
&= \frac{1}{1-(\frac{1}{a}-\Delta^\star)^2} + \frac{1}{1-(\frac{1}{a}-\Delta)^2}\\
&\leq \frac{1}{1-(\frac{1}{a}-\Delta)^2} + \frac{1}{1-(\frac{1}{a}-\Delta)^2}\\
&= \frac{2}{1-(\frac{1}{a}-\Delta)^2} && (10.4)
\end{aligned}$$

(A): $0 \leq \Delta \leq \frac{1}{a}$

(B): $0 \leq \Delta \leq \Delta^\star \leq \frac{1}{a}$

$$\begin{aligned}
P &\leq k^2(\frac{(2ak-k^2)\Sigma_E+1}{1-(a-k)^2}-\Sigma_E) \\
&\leq k^2(\frac{(2ak-k^2)\Sigma_E+1}{1-(a-k)^2}) \\
&\overset{(A)}{\leq} k^2\frac{2}{1-(\frac{1}{a}-\Delta)^2} \\
&= (a-\frac{1}{a}+\Delta)^2\frac{2}{1-(\frac{1}{a}-\Delta)^2} \\
&\overset{(B)}{\leq} (a+\Delta a-\frac{1}{a}+\Delta)^2\frac{2}{1-(\frac{1}{a}-\Delta)^2} \\
&= ((a+1)(1-\frac{1}{a}+\Delta))^2\frac{2}{1-(\frac{1}{a}-\Delta)^2} \\
&= \frac{2(a+1)^2(1-\frac{1}{a}+\Delta)^2}{1-(\frac{1}{a}-\Delta)^2} \\
&= \frac{2(a+1)^2(1-\frac{1}{a}+\Delta)}{1+\frac{1}{a}-\Delta} \\
&\overset{(C)}{\leq} 2(a+1)^2(1-\frac{1}{a}+\Delta)(1+\frac{1}{a}-\Delta) \\
&= 2(a+1)^2(1-(\frac{1}{a}-\Delta)^2)
\end{aligned}$$

(A): This comes from the comparison of (10.3) and (10.4).

(B): Since $\Delta \geq 0$, $a > 1$, we have $a - \frac{1}{a} + \Delta > 0$. Moreover, $\Delta a \geq 0$.

(C): $0 \leq \Delta \leq \frac{1}{a}$, $(1 + \frac{1}{a} - \Delta) \geq 1$.

Therefore, by putting $t = 2(a+1)^2(1-(\frac{1}{a}-\Delta)^2)$ we can conclude

$$(D_{\sigma_v}(P), P) \leq (\frac{4(a+1)^2}{t}, t).$$

Since $\Delta \in [0, \Delta^\star]$, we have $t \in [2(a+1)^2(1-(\frac{1}{a})^2), 2(a+1)^2(1-(\frac{1}{a}-\Delta^\star)^2)]$. Moreover, by the definition of $\Delta^\star$, it is equivalent to $t \in [2(a+1)^2(1-(\frac{1}{a})^2), \frac{2(a+1)^2}{\max(1,(a^2+1)\Sigma_E)}]$.

This finishes the proof of (ii).

When $1 < |a| \leq 2.5$, (i) is upper bounded as

$$\begin{aligned}
(D_{\sigma_v}(P), P) &\leq ((a^2+1)\Sigma_E + \frac{a^2}{a^2-1}, (a^2-1)^2\Sigma_E + (a^2-1)) \\
&\leq (7.25\Sigma_E + \frac{6.25}{a^2-1}, (a^2-1)^2\Sigma_E + (a^2-1)).
\end{aligned}$$

Thus, we get (i').

When $1 < |a| \leq 2.5$, (ii) is also upper bounded as

$$\begin{aligned}
(D_{\sigma_v}(P), P) &\leq (\frac{4(|a|+1)^2}{t}, t) \\
&\leq (\frac{49}{t}, t).
\end{aligned}$$

Moreover, for $1 < |a| \leq 2.5$

$$\begin{aligned}
&2(|a|+1)^2(1-(\frac{1}{a})^2) \leq t \leq \frac{2(|a|+1)^2}{\max(1,(a^2+1)\Sigma_E)} \\
&(\Leftrightarrow) 2(1+\frac{1}{|a|})^2(a^2-1) \leq t \leq \frac{2(|a|+1)^2}{\max(1,(a^2+1)\Sigma_E)} \\
&(\Rightarrow) 8(a^2-1) \leq t \leq \frac{8}{\max(1,7.25\Sigma_E)}.
\end{aligned}$$

Therefore, we get (ii'). □

Proof of Corollary 5.3 of Page 235. For simplicity, we will only proof for the case when $0 \leq a < 1$. The proof for $-1 < a \leq 0$ follows similarly by replacing a with $-a$.

First part of the lemma easily follows by putting $k = 0$ in Lemma 5.1.

Let's prove the second part of the lemma. Since the second part of the lemma assumes $\Sigma_E \leq \frac{1}{1-a^2}$, there always exists $k^\star \in [0, a]$ such that $\max(1, \Sigma_E) = \frac{1}{1-(a-k^\star)^2}$.

Since $0 \leq k^\star \leq a$ and $0 \leq a < 1$, for all $k \in [0, k^\star]$ we have $|a-k| < 1$. Thus, by Lemma 5.1, for all $k \in [0, k^\star]$ we have the following upper bounds on $D_{\sigma_v}(P), P$.

$$\begin{aligned}
D_{\sigma_v}(P) &\leq \frac{(2ak-k^2)\Sigma_E+1}{1-(a-k)^2} && (10.5) \\
&\overset{(A)}{\leq} \frac{(1-a^2+2ak-k^2)\Sigma_E+1}{1-(a-k)^2} \\
&= \frac{(1-(a-k)^2)\Sigma_E+1}{1-(a-k)^2} \\
&= \Sigma_E + \frac{1}{1-(a-k)^2} && (10.6) \\
&\overset{(B)}{\leq} \frac{1}{1-(a-k^\star)^2} + \frac{1}{1-(a-k)^2} \\
&\overset{(C)}{\leq} \frac{2}{1-(a-k)^2}
\end{aligned}$$

(A): $0 \leq a < 1$, $0 < k \leq a$, $\Sigma_E \geq 0$.
(B): $\max(1, \Sigma_E) = \frac{1}{1-(a-k^\star)^2}$.
(C): $0 \leq k \leq k^\star \leq a$.

$$
\begin{aligned}
P &\le k^2\Big(\frac{(2ak-k^2)\Sigma_E+1}{1-(a-k)^2}-\Sigma_E\Big)\\
&\overset{(A')}{\le} k^2\Big(\Sigma_E+\frac{1}{1-(a-k)^2}-\Sigma_E\Big)\\
&= \frac{k^2}{1-(a-k)^2}\\
&\overset{(B')}{\le} \frac{(1-a+k)^2}{1-(a-k)^2}\\
&= \frac{1-a+k}{1+a-k}\\
&\overset{(C')}{\le} (1-a+k)(1+a-k)\\
&= 1-(a-k)^2
\end{aligned}
$$

(A'): (10.5) $\le$ (10.6).

(B'): $0 \le a < 1$ and $0 < k \le a$.

(C'): $0 \le a < 1$ and $0 < k \le a$.

Let's put $t = 1-(a-k)^2$. Then, we have $(D_{\sigma_1}(P), P) \le (\frac{2}{t}, t)$. Moreover, since $0 \le k \le k^\star \le a$, $t \in [1-a^2, 1-(a-k^\star)^2]$. Furthermore, since t, $t \in [1-a^4 52, \frac{1}{\max(1,\Sigma_E)}]$. This finishes the proof of the lemma. □

10.2 Proof of Corollary 5.4 and Proposition 5.1

Proof of Corollary 5.4 of page 246. For simplicity, we first prove for the case when $1 < a \le 2.5$. The proof for the case when $-2.5 \le a < -1$ follows similarly.

First, let's upper bound Σ_1 and Σ_2 of (5.27) and (5.28). When $|(a^2-1)\sigma_{v1}^2 - 1| \ge |2a\sigma_{v1}|$, we have

$$
\begin{aligned}
\Sigma_1 &\le \frac{(a^2-1)\sigma_{v1}^2 - 1 + \sqrt{2((a^2-1)\sigma_{v1}^2-1)^2}}{2a^2}\\
&\le \frac{(1+\sqrt{2})|(a^2-1)\sigma_{v1}^2-1|}{2a^2}\\
&\le \frac{(1+\sqrt{2})\max(1,(a^2-1)\sigma_{v1}^2)}{2a^2} \qquad (10.7)
\end{aligned}
$$

When $|(a^2-1)\sigma_{v1}^2 - 1| \le |2a\sigma_{v1}|$, we have

$$
\begin{aligned}
\Sigma_1 &\le \frac{|2a\sigma_{v1}| + \sqrt{(2a\sigma_{v1})^2+4a^2\sigma_{v1}^2}}{2a^2}\\
&= \frac{(1+\sqrt{2})2a\sigma_{v1}}{2a^2} \qquad (10.8)
\end{aligned}
$$

Therefore, by (10.7) and (10.8), we can conclude

$$
\Sigma_1 \le \frac{(1+\sqrt{2})\max(1,(a^2-1)\sigma_{v1}^2, 2a\sigma_{v1})}{2a^2}. \qquad (10.9)
$$

Likewise, we also have

$$\Sigma_2 \leq \frac{(1+\sqrt{2})\max(1,(a^2-1)\sigma_{v2}^2, 2a\sigma_{v2})}{2a^2}.$$

We also have for all $k \geq 3$

$$\frac{a^2(1-a^{-2(k-1)})}{1-a^{-2(k-2)}} = \frac{a^{2(k-1)}-1}{a^{2(k-2)}-1} = \frac{(a-1)(1+\cdots+a^{(2k-4)}+a^{(2k-3)})}{(a-1)(1+\cdots+a^{(2k-5)})} \tag{10.10}$$

$$= \frac{1+\cdots+a^{(2k-4)}+a^{(2k-3)}}{1+\cdots+a^{(2k-5)}}$$

$$= 1+\frac{a^{(2k-4)}+a^{(2k-3)}}{1+\cdots+a^{(2k-5)}}$$

$$\overset{(A)}{\leq} 1+\frac{a^{(2k-4)}+a^{(2k-3)}}{a^{(2k-6)}+a^{(2k-5)}}$$

$$= 1+a^2 \leq 1+2.5^2 = 7.25. \tag{10.11}$$

(A): This comes from $k \geq 3$.

Then, let's prove the statements of the lemma.

Proof of (a):

Since $\Sigma_1 \geq 150$ and $\Sigma_2 \geq 150$, there exist $k_1 \geq 3$ and $k_2 \geq 3$ such that

$$\frac{a^{2(k_1-2)}-1}{1-a^{-2}} \leq \frac{\Sigma_1}{24} < \frac{a^{2(k_1-1)}-1}{1-a^{-2}} \tag{10.12}$$

$$\frac{a^{2(k_2-2)}-1}{1-a^{-2}} \leq \frac{\Sigma_2}{24} < \frac{a^{2(k_2-1)}-1}{1-a^{-2}}$$

We will evaluate Lemma 5.3 with these k_1 and k_2, and increase k arbitrary large.

Moreover, since $\Sigma_1 \geq 150$ implies $\sigma_{v1} \geq 1$, (10.9) further reduces to

$$\Sigma_1 \leq \frac{(1+\sqrt{2})\max((a^2-1)\sigma_{v1}^2, 2a\sigma_{v1})}{2a^2}. \tag{10.13}$$

Let's upper bound I of Lemma 5.3. First, we have

$$\frac{a^{2(k_1-2)}(1-a^{-2(k_1-1)})^2}{(1-a^{-2})^2}$$

$$\overset{(A)}{\leq} \frac{a^{2(k_1-2)}(1-a^{-2(k_1-1)})}{(1-a^{-2})^2}$$

$$\overset{(B)}{\leq} \frac{a^{2(k_1-2)}(7.25a^{-2}(1-a^{-2(k_1-2)}))}{(1-a^{-2})^2}$$

$$= 7.25a^{-2}(\frac{a^{2(k_1-2)}-1}{1-a^{-2}})\frac{1}{1-a^{-2}}$$

$$\leq \frac{7.25\Sigma_1}{24}\frac{a^{-2}}{1-a^{-2}} = \frac{7.25\Sigma_1}{24}\frac{1}{a^2-1}. \tag{10.14}$$

(A): For $k_1 \geq 3$, $1-a^{-2(k_1-1)} \leq 1$.

(B): By comparing (10.10) and (10.11), we get $7.25a^{-2}(1-a^{-2(k_1-2)}) \geq (1-a^{-2(k_1-1)})$.

(C): This comes from (10.12).

Moreover, we also have

$$\begin{aligned}
&\frac{a^{2(k_1-2)}(1-a^{-2(k_1-1)})^2}{(1-a^{-2})^2} \\
&\overset{(A)}{\leq} \frac{a^{2(k_1-2)}(7.25a^{-2}(1-a^{-2(k_1-2)}))^2}{(1-a^{-2})^2} \\
&\overset{(B)}{\leq} 7.25^2(\frac{a^{2(k_1-2)}(1-a^{-2(k_1-2)})}{1-a^{-2}})^2 \\
&\overset{(C)}{\leq} (\frac{7.25\Sigma_1}{24})^2. \qquad (10.15)
\end{aligned}$$

(A): By comparing (10.10) and (10.11), we get $7.25a^{-2}(1-a^{-2(k_1-2)}) \geq (1-a^{-2(k_1-1)})$.

(B): $a > 1$ and $k_1 \geq 3$.

(C): This comes from (10.12).

By merging the results so far, we can conclude

$$\begin{aligned}
&\frac{a^{2(k_1-2)}(1-a^{-2(k_1-1)})^2}{(1-a^{-2})^2} \qquad (10.16) \\
&\overset{(A)}{\leq} \min(\frac{7.25\Sigma_1}{24}\frac{1}{a^2-1}, (\frac{7.25\Sigma_1}{24})^2) \qquad (10.17) \\
&\overset{(B)}{\leq} \max(\frac{7.25}{24}\frac{1}{a^2-1}\frac{(1+\sqrt{2})(a^2-1)\sigma_{v1}^2}{2a^2}, (\frac{7.25}{24})^2(\frac{(1+\sqrt{2})2a\sigma_{v1}}{2a^2})^2) \\
&= \max(\frac{7.25}{24}\frac{1+\sqrt{2}}{2a^2}, (\frac{7.25}{24})^2(\frac{1+\sqrt{2}}{a})^2)\sigma_{v1}^2 \\
&\overset{(C)}{\leq} \max(\frac{7.25}{24}\frac{1+\sqrt{2}}{2}, (\frac{7.25}{24})^2(1+\sqrt{2})^2)\sigma_{v1}^2 \\
&\leq 0.5319\sigma_{v1}^2 \qquad (10.18)
\end{aligned}$$

(A): This comes from (10.14) and (10.15).

(B): When $(a^2-1)\sigma_{v1}^2 \geq 2a\sigma_{v1}$, by (10.13) we have $\Sigma_1 \leq \frac{(1+\sqrt{2})(a^2-1)\sigma_{v1}^2}{2a^2}$. Thus, by plugging it into (10.17), we get

$$(10.16) \leq \frac{7.25}{24}\frac{1}{a^2-1}\frac{(1+\sqrt{2})(a^2-1)\sigma_{v1}^2}{2a^2}.$$

Likewise, when $(a^2-1)\sigma_{v1}^2 \leq 2a\sigma_{v1}$, by (10.13) we have $\Sigma_1 \leq \frac{(1+\sqrt{2})2a\sigma_{v1}}{2a^2}$. Therefore, by plugging it into (10.17), we get

$$(10.16) \leq (\frac{7.25}{24})^2(\frac{(1+\sqrt{2})2a\sigma_{v1}}{2a^2})^2.$$

(C): Because $a > 1$.

In the same ways, we can also prove that

$$\frac{a^{2(k_2-2)}(1-a^{-2(k_2-1)})^2}{(1-a^{-2})^2} \leq 0.5319\sigma_{v2}^2. \qquad (10.19)$$

Therefore, by plugging (10.18) and (10.19) into I of Lemma 5.3, we can upper bound I by

$$\begin{aligned} I &\leq (k_1 - 1)\log(1 + \frac{1}{k_1 - 1}0.5319) \\ &\leq \log e^{0.5319}. \end{aligned} \tag{10.20}$$

Let's upper bound $I'(\widetilde{P_1})$. First, we have

$$\begin{aligned} &2a^{2(k_2-1-k)}\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}}\Sigma + 2a^{2(k_2-1-k_1)}\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}}\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}} \\ &\overset{(A)}{\leq} 2a^{2(k_2-1-k)}\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}}\frac{a^{2(k-1)}(1-a^{-2(k_1-1)})}{1-a^{-2}} \\ &+ 2a^{2(k_2-1-k_1)}\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}}\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}} \\ &= 2a^{2(k_2-2)}(\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}})(\frac{(1-a^{-2(k_1-1)})}{1-a^{-2}} + a^{2(-k_1+1)}\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}}) \\ &= 2a^{2(k_2-2)}(\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}})(\frac{1-a^{-2(k_1-1)}+a^{-2(k_1-1)}-a^{-2(k_2-1)}}{1-a^{-2}}) \\ &\overset{(B)}{\leq} 2a^{2(k_2-2)}(\frac{1-a^{-2(k_2-1)}}{1-a^{-2}})^2 \\ &\overset{(C)}{\leq} 2\cdot 0.5319\sigma_{v2}^2. \end{aligned} \tag{10.21}$$

(A): Since $I \geq 0$, $\Sigma \leq a^{2(k-1)}\frac{1-a^{-2(k_1-1)}}{1-a^{-2}}$.

(B): $k_1 \geq 1$.

(C): It comes from (10.19).

We also have

$$\begin{aligned} &2a^{2(k_2-k_1-2)}\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}}\frac{(1-a^{-(k_2-1-k_1)})(1-a^{-(k-k_1)})}{(1-a^{-1})^2}\widetilde{P_1} \\ &\overset{(A)}{\leq} 2a^{2(k_2-k_1-2)}\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}}\frac{(1-a^{-(k_2-1-k_1)})(1-a^{-(k-k_1)})}{(1-a^{-1})^2}\frac{24(a^2-1)^2}{40000}\frac{a^{2(k_1-1)}-1}{1-a^{-2}} \\ &\overset{(B)}{\leq} 2a^{2(k_2-2)}(\frac{1-a^{-2(k_2-1)}}{1-a^{-2}})^2\frac{24a^{-2}}{40000}\frac{(a^2-1)^2}{(1-a^{-1})^2} \\ &= 2a^{2(k_2-2)}(\frac{1-a^{-2(k_2-1)}}{1-a^{-2}})^2\frac{24(a+1)^2}{40000} \\ &\leq \frac{48(2.5+1)^2}{40000}0.5319\sigma_{v2}^2 \\ &= 0.00781893\sigma_{v2}^2. \end{aligned} \tag{10.22}$$

(A): Since we have $\widetilde{P_1} \leq \frac{(a^2-1)^2\Sigma_1}{40000}$ and $\Sigma_1 \leq 24\frac{a^{2(k_1-1)}-1}{1-a^{-2}}$ by (10.12).

(B): Since $k_2 - 1 \geq k_2 - k_1$ and $2(k_2-1) \geq (k_2 - 1 - k_1)$.

(C): By (10.19) and $0 \leq a \leq 2.5$.

Therefore, by (10.21) and (10.22), we can bound $I'(\widetilde{P_1})$ of Lemma 5.3 by

$$\begin{aligned} I'(\widetilde{P_1}) &\leq \frac{k_2 - k_1}{2}\log(1 + \frac{1}{k_2 - k_1}(2 \cdot 0.5319 + 0.00781893)) \\ &\leq \frac{k_2 - k_1}{2}\log(1 + \frac{1}{k_2 - k_1}(1.07161893)) \\ &\leq \frac{1}{2}\log e^{1.0717}. \end{aligned} \tag{10.23}$$

Moreover, we have

$$\begin{aligned} & a^{2(k-k_1-1)}\frac{(1-a^{-(k-k_1)})^2}{(1-a^{-1})^2}\widetilde{P_1} \\ &\overset{(A)}{\leq} a^{2(k-k_1-1)}\frac{(1-a^{-(k-k_1)})^2}{(1-a^{-1})^2}\frac{24(a^2-1)^2}{40000}\frac{a^{2(k_1-1)}-1}{1-a^{-2}} \\ &= a^{2(k-k_1-1)}\frac{1-2a^{-(k-k_1)}+a^{-2(k-k_1)}}{(1-a^{-1})^2}\frac{24(a^2-1)^2}{40000}\frac{a^{2(k_1-1)}-1}{1-a^{-2}} \\ &\overset{(B)}{\leq} a^{2(k-k_1-1)}\frac{1-a^{-2(k-k_1)}}{(1-a^{-1})^2}\frac{24(a^2-1)^2}{40000}\frac{a^{2(k_1-1)}-1}{1-a^{-2}} \\ &= \frac{a^{2(k-2)}(1-a^{-2(k-k_1)})(1-a^{-2(k_1-1)})}{(1-a^{-2})} \cdot \frac{24(a^2-1)^2}{40000(1-a^{-1})^2} \\ &\leq \frac{a^{2(k-2)}(1-a^{-2(k-1)})}{(1-a^{-2})} \cdot \frac{24(a^2-1)^2}{40000(1-a^{-1})^2} \\ &= \frac{a^{2(k-1)}(1-a^{-2(k-1)})}{(1-a^{-2})} \cdot \frac{24(a+1)^2}{40000} \\ &\overset{(C)}{\leq} \frac{a^{2(k-1)}(1-a^{-2(k-1)})}{(1-a^{-2})} \cdot \frac{24(2.5+1)^2}{40000} \\ &= \frac{a^{2(k-1)}(1-a^{-2(k-1)})}{(1-a^{-2})} \cdot \frac{147}{20000} \end{aligned} \tag{10.24}$$

(A): By (10.12) and $\widetilde{P_1} \leq \frac{(a^2-1)^2\Sigma_1}{40000}$.

(B): Since $k \geq k_1$.

(C): Since $1 \leq a \leq 2.5$.

Likewise, we can also prove that

$$a^{2(k-k_2-1)}\frac{(1-a^{-(k-k_2)})^2}{(1-a^{-1})^2}\widetilde{P_2} \leq \frac{a^{2(k-1)}(1-a^{-2k-1)})}{(1-a^{-2})} \cdot \frac{147}{20000}. \tag{10.25}$$

Finally, by plugging (10.20), (10.23), (10.24), (10.25) into Lemma 5.3 we have

$$\begin{aligned}
&D_L(\widetilde{P_1},\widetilde{P_2})\\
&\geq (\sqrt{\frac{a^{2(k-1)}\frac{1-a^{-2(k_1-1)}}{1-a^{-2}}+a^{2(k-k_1)}\frac{1-a^{-2(k_2-k_1)}}{1-a^{-2}}+a^{2(k-k_2)}\frac{1-a^{-2(k-k_2)}}{1-a^{-2}}}{2^{2(I+I'(\widetilde{P_1}))}}}\\
&-\sqrt{a^{2(k-k_1-1)}\frac{(1-a^{-(k-k_1)})^2}{(1-a^{-1})^2}\widetilde{P_1}}-\sqrt{a^{2(k-k_2-1)}\frac{(1-a^{-(k-k_2)})^2}{(1-a^{-1})^2}\widetilde{P_2}})_+^2+1\\
&\geq \frac{a^{2(k-1)}(1-a^{-2(k-1)})}{1-a^{-2}}(\sqrt{\frac{1}{e^{2\cdot 0.5319+1.0717}}}-\sqrt{\frac{147}{20000}}-\sqrt{\frac{147}{20000}})_+^2+1\\
&\geq \frac{a^{2(k-1)}(1-a^{-2(k-1)})}{1-a^{-2}}0.02969+1.
\end{aligned}$$

Therefore, by choosing k arbitrary large, we have $D_L(\widetilde{P_1},\widetilde{P_2})=\infty$.

Proof of (b):

Like (a), since $\Sigma_1\geq 150$ and $\Sigma_2\geq 150$, there exist $k_1\geq 3$ and $k_2\geq 3$ such that

$$\begin{aligned}
\frac{a^{2(k_1-2)}-1}{1-a^{-2}}&\leq\frac{\Sigma_1}{24}<\frac{a^{2(k_1-1)}-1}{1-a^{-2}},\\
\frac{a^{2(k_2-2)}-1}{1-a^{-2}}&\leq\frac{\Sigma_2}{24}<\frac{a^{2(k_2-1)}-1}{1-a^{-2}}.
\end{aligned}$$

We put the parameters of Lemma 5.3 as such k_1, k_2 and $k=k_2$. Then, the lower bound of Lemma 5.3 reduces to

$$D_L(\widetilde{P_1},\widetilde{P_2})\geq(\sqrt{\frac{\Sigma+a^{2(k-k_1)}\frac{1-a^{-2(k-k_1)}}{1-a^{-2}}}{2^{2I'(\widetilde{P_1})}}}-\sqrt{a^{2(k-k_1-1)}\frac{(1-a^{-(k-k_1)})^2}{(1-a^{-1})^2}\widetilde{P_1}})_+^2+1.$$

Since we choose k_1 and k_2 in the same way as (a) and have the same bound on $\widetilde{P_1}$, we still have (10.20), (10.23), (10.24) which are

$$\begin{aligned}
&I\leq\log e^{0.5319},\\
&I'(\widetilde{P_1})\leq\frac{1}{2}\log e^{1.0717},\\
&a^{2(k-k_1-1)}\frac{(1-a^{-(k-k_1)})^2}{(1-a^{-1})^2}\widetilde{P_1}\leq\frac{a^{2(k-1)}(1-a^{-2(k-1)})}{(1-a^{-2})}\cdot\frac{147}{20000}.
\end{aligned}$$

Therefore, we can conclude

$$\begin{aligned}
&D_L(\widetilde{P_1},\widetilde{P_2})\\
&\geq(\sqrt{\frac{a^{2(k-1)}\frac{1-a^{-2(k_1-1)}}{1-a^{-2}}+a^{2(k-k_1)}\frac{1-a^{-2(k-k_1)}}{1-a^{-2}}}{2^{2(I+I'(P_1))}}}-\sqrt{a^{2(k-k_1-1)}\frac{(1-a^{-(k-k_1)})^2}{(1-a^{-1})^2}P_1})_+^2+1\\
&\geq\frac{a^{2(k-1)}(1-a^{-2(k-1)})}{1-a^{-2}}(\sqrt{\frac{1}{e^{2\cdot 0.5319+1.0717}}}-\sqrt{\frac{147}{20000}})_+^2+1\\
&\geq\frac{\Sigma_2}{24}(\sqrt{\frac{1}{e^{2\cdot 0.5319+1.0717}}}-\sqrt{\frac{147}{20000}})_+^2+1\\
&\geq 0.002774\Sigma_2+1.
\end{aligned}$$

Proof of (c):

We will put $k_1 = k_2 = 1$ in Lemma 5.3 and increase k arbitrary large. First, the lower bound in Lemma 5.3 reduces to

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq (\sqrt{a^{2(k-1)}\frac{1-a^{-2(k-1)}}{1-a^{-2}}} - \sqrt{\frac{a^{2(k-2)}(1-a^{-(k-1)})^2}{(1-a^{-1})^2}\widetilde{P_1}} - \sqrt{\frac{a^{2(k-2)}(1-a^{-(k-1)})^2}{(1-a^{-1})^2}\widetilde{P_2}})_+^2 + 1. \tag{10.26}$$

Here, we have

$$\begin{aligned} &\frac{a^{2(k-2)}(1-a^{-(k-1)})^2}{(1-a^{-1})^2}\widetilde{P_1} \\ &\leq \frac{a^{2(k-2)}(1-2a^{-(k-1)}+a^{-2(k-1)})}{(1-a^{-1})^2}\frac{1}{20}(a^2-1) \\ &\overset{(A)}{\leq} \frac{a^{2(k-1)}(1-a^{-2(k-1)})}{(1-a^{-2})}(1+a^{-1})^2\frac{1}{20} \\ &\overset{(B)}{\leq} \frac{1}{5}\frac{a^{2(k-1)}(1-a^{-2(k-1)})}{(1-a^{-2})} \end{aligned} \tag{10.27}$$

(A): Since $k \geq 1$.

(B): Since $a \geq 1$.

Likewise, we can also prove that

$$\frac{a^{2(k-2)}(1-a^{-(k-1)})^2}{(1-a^{-1})^2}\widetilde{P_2} \leq \frac{1}{5}\frac{a^{2(k-1)}(1-a^{-2(k-1)})}{(1-a^{-2})} \tag{10.28}$$

Finally, by plugging (10.27), (10.28) into (10.26), we get

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq a^{2(k-1)}\frac{1-a^{-2(k-1)}}{1-a^{-2}}(1-\sqrt{\frac{1}{5}}-\sqrt{\frac{1}{5}})_+^2 + 1.$$

Therefore, by choosing k arbitrary large, we have $D_L(\widetilde{P_1}, \widetilde{P_2}) = \infty$.

Proof of (d):

Let $k_1 = k_2 = 1$ and $P = \max(\widetilde{P_1}, \widetilde{P_2})$. Since $P \leq \frac{1}{75}$, we can find $k \geq 2$ such that

$$\frac{a^{(k-1)}-1}{1-a^{-1}} \leq \frac{1}{30P} < \frac{a^k-1}{1-a^{-1}} \tag{10.29}$$

By setting the parameters of Lemma 5.3 to such k_1, k_2, k, the lower bound in Lemma 5.3 reduces to

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq (\sqrt{a^{2(k-1)}\frac{1-a^{-2(k-1)}}{1-a^{-2}}} - \sqrt{\frac{a^{2(k-2)}(1-a^{-(k-1)})^2}{(1-a^{-1})^2}\widetilde{P_1}} - \sqrt{\frac{a^{2(k-2)}(1-a^{-(k-1)})^2}{(1-a^{-1})^2}\widetilde{P_2}})_+^2 + 1. \tag{10.30}$$

The first term of (10.30) is lower bounded as follows:

$$\begin{aligned} &a^{2(k-1)}\frac{1-a^{-2(k-1)}}{1-a^{-2}} \\ &\overset{(A)}{\geq} \frac{a^k-1}{1-a^{-1}}\frac{1}{1+a^{-1}} \\ &\overset{(B)}{\geq} \frac{1}{60P} \end{aligned} \tag{10.31}$$

(A): $k \geq 2$.

(B): (10.29) and $a \geq 1$.

The second term of (10.30) is upper bounded as follows:

$$\frac{a^{2(k-2)}(1-a^{-(k-1)})^2}{(1-a^{-1})^2}\widetilde{P_1} \overset{(A)}{\leq} \frac{(a^{k-1}-1)^2}{(1-a^{-1})^2}\widetilde{P_1} \overset{(B)}{\leq} \frac{1}{900P^2}\widetilde{P_1} \leq \frac{1}{900P} \tag{10.32}$$

(A): $a \geq 1$.

(B): (10.29).

Likewise, the third term of (10.30) is upper bounded as

$$\frac{a^{2(k-2)}(1-a^{-(k-1)})^2}{(1-a^{-1})^2}\widetilde{P_1} \leq \frac{1}{900P}. \tag{10.33}$$

Therefore, by plugging (10.31), (10.32), (10.33) into (10.30), we conclude

$$\begin{aligned} D_L(\widetilde{P_1}, \widetilde{P_2}) &\geq \frac{1}{P}\Big(\sqrt{\frac{1}{60}} - \sqrt{\frac{1}{900}} - \sqrt{\frac{1}{900}}\Big)^2 + 1 \\ &\geq 0.00389\frac{1}{P} + 1 \end{aligned}$$

Proof of (e):

Since $\Sigma_2 \geq 150$, we can find $k \geq 3$ such that

$$\frac{a^{2(k-2)}-1}{1-a^{-2}} \leq \frac{\Sigma_2}{24} < \frac{a^{2(k-1)}-1}{1-a^{-2}}. \tag{10.34}$$

Let $k_2 = k$ and $k_1 = 1$. By putting the parameters of Lemma 5.3 with these parameters, the lower bound of Lemma 5.3 reduces to

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq \Bigg(\sqrt{\frac{a^{2(k-1)}\frac{1-a^{-2(k-1)}}{1-a^{-2}}}{2^{2I'(\widetilde{P_1})}}} - \sqrt{\frac{a^{2(k-2)}(1-a^{-(k-1)})^2}{(1-a^{-1})^2}\widetilde{P_1}}\Bigg)_+^2 + 1. \tag{10.35}$$

We will upper bound $I'(\widetilde{P_1})$. First, since we chose k in the same ways as k_2 of (a), (10.21) still holds, i.e.

$$2a^{2(k-2)}\frac{1-a^{-2(k-1)}}{1-a^{-2}}\Sigma + 2a^{2(k-2)}\frac{1-a^{-2(k-1)}}{1-a^{-2}}\frac{1-a^{-2(k-1)}}{1-a^{-2}} \leq 2\cdot 0.5319\sigma_{v2}^2. \tag{10.36}$$

Moreover, we have

$$\begin{aligned} & 2a^{2(k-3)}\frac{1-a^{-2(k-1)}}{1-a^{-2}}\frac{(1-a^{-(k-2)})(1-a^{-(k-1)})}{(1-a^{-1})^2}\widetilde{P_1} \\ & \overset{(A)}{\leq} 2a^{2(k-3)}\frac{1-a^{-2(k-1)}}{1-a^{-2}}\frac{(1-a^{-(k-2)})(1-a^{-(k-1)})}{(1-a^{-1})^2}\frac{1}{24}\frac{1-a^{-2}}{a^{2(k-2)}-1} \\ & \overset{(B)}{\leq} \frac{1}{12}\frac{a^{-4}(a^{2(k-1)}-1)}{a^{2(k-2)}-1}\frac{(1-a^{-(k-1)})^2}{(1-a^{-1})^2} \\ & \overset{(C)}{\leq} \frac{1}{12}(1+a^{-1})^2a^{-2}\frac{a^{2(k-2)}(1-a^{-2(k-1)})^2}{(1-a^{-2})^2} \\ & \overset{(D)}{\leq} \frac{1}{3}0.5319\sigma_{v2}^2 \end{aligned} \tag{10.37}$$

(A): $\widetilde{P_1} \leq \frac{1}{\Sigma_2} \leq \frac{1}{24}\frac{1-a^{-2}}{a^{2(k-2)}-1}$.

(B): $1 - a^{-(k-2)} \leq 1 - a^{-(k-1)}$.

(C): Since $k \geq 3$ and $1 \leq a \leq 2.5$, we have

$$\begin{aligned}
&(a^{2(k-1)} - 1) \leq (a^{4(k-2)} - 1) \\
(\Leftrightarrow)&(a^{2(k-1)} - 1) \leq (a^{2(k-2)} - 1)(a^{2(k-2)} + 1) \\
(\Rightarrow)&(a^{2(k-1)} - 1) \leq (a^{2(k-2)} - 1)(a^{k-1} + 1)^2 \\
(\Leftrightarrow)&(a^{2(k-1)} - 1) \leq a^{2(k-1)}(a^{2(k-2)} - 1)(1 + a^{-(k-1)})^2 \\
(\Leftrightarrow)&\frac{a^{-4}(a^{2(k-1)} - 1)(1 - a^{-(k-1)})^2}{(a^{2(k-2)} - 1)(1 - a^{-1})^2} \leq (1 + a^{-1})^2 a^{-2}\frac{a^{2(k-2)}(1 - a^{-2(k-1)})^2}{(1 - a^{-2})^2}.
\end{aligned}$$

(D): This comes from $1 \leq a \leq 2.5$ and (10.19).

Therefore, by (10.36) and (10.37), we have

$$\begin{aligned}
I'(\widetilde{P_1}) &\leq \frac{1}{2}\log(1 + \frac{1}{k-1}(2 \cdot 0.5319 + \frac{1}{3}0.5319)^{k-1} \\
&\leq \frac{1}{2}\log(1 + \frac{1}{k-1}(1.2411))^{k-1} \\
&\leq \frac{1}{2}\log e^{1.2411}. \qquad (10.38)
\end{aligned}$$

We also have

$$\begin{aligned}
&\frac{a^{2(k-2)}(1 - a^{-(k-1)})^2}{(1 - a^{-1})^2}\widetilde{P_1} \\
&= \frac{a^{-2}(a^{k-1} - 1)^2}{(1 - a^{-1})^2}\widetilde{P_1} \\
&\overset{(A)}{\leq} \frac{a^{-2}(a^{2(k-2)} - 1)^2}{(1 - a^{-2})^2}(1 + a^{-1})^2\widetilde{P_1} \\
&\overset{(B)}{\leq} (\frac{\Sigma_2}{24})^2 4\widetilde{P_1} \\
&\overset{(C)}{\leq} \frac{\Sigma_2}{144} \qquad (10.39)
\end{aligned}$$

(A): This comes from $2(k-2) \geq (k-1)$.

(B): By (10.34) and $1 \leq a \leq 2.5$.

(C): Since $P_1 \leq \frac{1}{\Sigma_2}$.

Therefore, by plugging (10.34), (10.38), (10.39) into (10.35), we get

$$\begin{aligned}
D_L(\widetilde{P_1}, \widetilde{P_2}) &\geq (\sqrt{\frac{a^{2(k-1)}\frac{1-a^{-2(k-1)}}{1-a^{-2}}}{2^{2I'(\widetilde{P_1})}}} - \sqrt{\frac{\Sigma_2}{144}})_+^2 + 1 \\
&\geq (\sqrt{\frac{\Sigma_2}{24 \cdot 2^{2I'(\widetilde{P_1})}}} - \sqrt{\frac{\Sigma_2}{144}})_+^2 + 1 \\
&\geq \Sigma_2(\sqrt{\frac{1}{24 \cdot e^{1.2411}}} - \sqrt{\frac{1}{144}})^2 + 1 \\
&\geq 0.0006976\Sigma_2 + 1.
\end{aligned}$$

Proof of (f):

Since $\widetilde{P_1} \leq \frac{1}{150}$, there exists $k \geq 3$ such that

$$\frac{a^{2(k-2)}-1}{1-a^{-2}} \leq \frac{1}{24\widetilde{P_1}} < \frac{a^{2(k-1)}-1}{1-a^{-2}}. \tag{10.40}$$

Let $k_2 = k$ and $k_1 = 1$. By putting the parameters of Lemma 5.3 as these parameters, the lower bound of Lemma 5.3 reduces to

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq (\sqrt{\frac{a^{2(k-1)}\frac{1-a^{-2(k-1)}}{1-a^{-2}}}{2^{2I'(\widetilde{P_1})}}} - \sqrt{\frac{a^{2(k-2)}(1-a^{-(k-1)})^2}{(1-a^{-1})^2}\widetilde{P_1}})_+^2 + 1 \tag{10.41}$$

We will upper bound $I'(\widetilde{P_1})$. Since we assumed $\frac{1}{\widetilde{P_1}} \leq \Sigma_2$, by (10.40) we have $\frac{a^{2(k-2)}-1}{1-a^{-2}} \leq \frac{\Sigma_2}{24}$. Therefore, (10.21) still holds and we have

$$2a^{2(k-2)}\frac{1-a^{-2(k-1)}}{1-a^{-2}}\Sigma + 2a^{2(k-2)}\frac{1-a^{-2(k-1)}}{1-a^{-2}}\frac{1-a^{-2(k-1)}}{1-a^{-2}} \leq 2 \cdot 0.5319\sigma_{v2}^2.$$

Since $\widetilde{P_1} \leq \frac{1}{24}\frac{1-a^{-2}}{a^{2(k-2)}-1}$, following the same process of (10.37) we have

$$\begin{aligned}&2a^{2(k-3)}\frac{1-a^{-2(k-1)}}{1-a^{-2}}\frac{(1-a^{-(k-2)})(1-a^{-(k-1)})}{(1-a^{-1})^2}\widetilde{P_1}\\&\leq \frac{1}{3}0.5319\sigma_{v2}^2.\end{aligned}$$

Therefore, $I'(\widetilde{P_1})$ is upper bounded by

$$\begin{aligned}I'(\widetilde{P_1}) &\leq \frac{1}{2}\log(1+\frac{1}{k-1}(2\cdot 0.5319 + \frac{1}{3}0.5319)^{k-1}\\&\leq \frac{1}{2}\log(1+\frac{1}{k-1}(1.2411))^{k-1}\\&\leq \frac{1}{2}\log e^{1.2411}\end{aligned} \tag{10.42}$$

Moreover, we also have

$$\begin{aligned}&\frac{a^{2(k-2)}(1-a^{-(k-1)})^2}{(1-a^{-1})^2}\widetilde{P_1}\\&= \frac{a^{-2}(a^{k-1}-1)^2}{(1-a^{-1})^2}\widetilde{P_1}\\&\overset{(A)}{\leq} \frac{a^{-2}(a^{2(k-2)}-1)^2}{(1-a^{-2})^2}(1+a^{-1})^2\widetilde{P_1}\\&\overset{(B)}{\leq} (\frac{1}{24\widetilde{P_1}})^2 4\widetilde{P_1} = \frac{1}{144\widetilde{P_1}}\end{aligned} \tag{10.43}$$

(A): This comes from $2(k-2) \geq (k-1)$.

(B): By (10.40) and $1 \leq a \leq 2.5$.

Therefore, by plugging (10.40), (10.42), (10.43) into (10.41), we have

$$\begin{aligned}
D_L(\widetilde{P_1}, \widetilde{P_2}) &\geq \left(\sqrt{\frac{a^{2(k-1)}\frac{1-a^{-2(k-1)}}{1-a^{-2}}}{2^{2I'(\widetilde{P_1})}}} - \sqrt{\frac{1}{144\widetilde{P_1}}}\right)_+^2 + 1 \\
&\geq \left(\sqrt{\frac{1}{24\widetilde{P_1}\cdot 2^{2I'(\widetilde{P_1})}}} - \sqrt{\frac{1}{144\widetilde{P_1}}}\right)_+^2 + 1 \\
&\geq \frac{1}{\widetilde{P_1}}\left(\sqrt{\frac{1}{24\cdot e^{1.2411}}} - \sqrt{\frac{1}{144}}\right)^2 + 1 \\
&\geq \frac{0.000697686...}{\widetilde{P_1}} + 1 \\
&\geq \frac{0.0006976}{\widetilde{P_1}} + 1.
\end{aligned}$$

Proof of (g):

Since $\Sigma_2 \geq 150$, we can find $k_2 \geq 3$ such that

$$\frac{a^{2(k_2-2)}-1}{1-a^{-2}} \leq \frac{\Sigma_2}{24} < \frac{a^{2(k_2-1)}-1}{1-a^{-2}}$$

Let $k_1 = 1$ and increase k arbitrary large. By plugging such parameters to Lemma 5.3, the lemma reduces to

$$\begin{aligned}
D_L(\widetilde{P_1}, \widetilde{P_2}) \geq &\left(\sqrt{\frac{a^{2(k-1)}\frac{1-a^{-2(k_2-1)}}{1-a^{-2}}}{2^{2I'(\widetilde{P_1})}} + a^{2(k-k_2)}\frac{1-a^{-2(k-k_2)}}{1-a^{-2}}}\right. \\
&\left. - \sqrt{a^{2(k-2)}\frac{(1-a^{-(k-1)})^2}{(1-a^{-1})^2}\widetilde{P_1}} - \sqrt{a^{2(k-k_2-1)}\frac{(1-a^{-(k-k_2)})^2}{(1-a^{-1})^2}\widetilde{P_2}}\right)_+^2 + 1. \quad (10.44)
\end{aligned}$$

We will first upper bound $I'(\widetilde{P_1})$. Following the same steps as (10.21), we get

$$2a^{2(k-2)}\frac{1-a^{-2(k-1)}}{1-a^{-2}}\Sigma + 2a^{2(k-2)}\frac{1-a^{-2(k-1)}}{1-a^{-2}}\frac{1-a^{-2(k-1)}}{1-a^{-2}} \leq 2\cdot 0.5319\sigma_{v2}^2.$$

We also have

$$\begin{aligned}
&a^{2(k_2-3)}\frac{1-a^{-2(k_2-1)}}{1-a^{-2}}\frac{(1-a^{-(k_2-2)})(1-a^{-(k-1)})}{(1-a^{-1})^2}\widetilde{P_1} \\
&\overset{(A)}{\leq} a^{2(k_2-3)}\frac{(1-a^{-2(k_2-1)})(1-a^{-(k_2-2)})}{(1-a^{-1})^2}\frac{\widetilde{P_1}}{1-a^{-2}} \\
&\overset{(B)}{\leq} a^{2(k_2-3)}\frac{(1-a^{-2(k_2-1)})^2}{(1-a^{-1})^2}\frac{1}{20}a^2 \\
&= \frac{a^{2(k_2-2)}(1-a^{-2(k_2-1)})^2}{(1-a^{-2})^2}\frac{1}{20}(1+a^{-1})^2 \\
&\overset{(C)}{\leq} \frac{a^{2(k_2-2)}(1-a^{-2(k_2-1)})^2}{(1-a^{-2})^2}\frac{1}{5} \\
&\overset{(D)}{\leq} \frac{0.5319}{5}\sigma_{v2}^2.
\end{aligned}$$

(A): Since $0 \leq 1 - a^{-(k-1)} \leq 1$.

(B): Since we assumed $\widetilde{P_1} \leq \frac{1}{20}(a^2 - 1)$.

(C): Since $1 \leq a \leq 2.5$.

(D): This follows from that (10.19) still holds.

Therefore, $I'(\widetilde{P_1})$ is upper bounded by

$$\begin{aligned} I'(\widetilde{P_1}) &\leq \frac{1}{2}\log(1 + \frac{(2+\frac{2}{5})0.5319}{k_2 - 1})^{k_2-1} \\ &\leq \frac{1}{2}\log(1 + \frac{1.27656}{k_2 - 1})^{k_2-1} \\ &\leq \frac{1}{2}\log e^{1.27656}. \end{aligned} \tag{10.45}$$

Following the same steps as (10.27), we still have

$$\frac{a^{2(k-2)}(1 - a^{-(k-1)})^2}{(1 - a^{-1})^2}\widetilde{P_1} \leq \frac{1}{5}\frac{a^{2(k-1)}(1 - a^{-2(k-1)})}{(1 - a^{-2})}. \tag{10.46}$$

Following the same steps as (10.24), we still have

$$a^{2(k-k_2-1)}\frac{(1 - a^{-2(k-k_2)})^2}{(1 - a^{-1})^2}\widetilde{P_2} \leq \frac{a^{2(k-1)}(1 - a^{-2(k-1)})}{(1 - a^{-2})}\frac{147}{20000} \tag{10.47}$$

Therefore, by plugging (10.45), (10.46), (10.47) into (10.44) we conclude

$$\begin{aligned} D_L(\widetilde{P_1}, \widetilde{P_2}) &\geq \frac{a^{2(k-1)}(1 - a^{-2(k-1)})}{(1 - a^{-2})}(\sqrt{\frac{1}{e^{1.27656}}} - \sqrt{\frac{1}{5}} - \sqrt{\frac{147}{20000}})^2 + 1 \\ &\geq 0.00002252\frac{a^{2(k-1)}(1 - a^{-2(k-1)})}{(1 - a^{-2})} + 1. \end{aligned}$$

Finally, by increasing k arbitrarily large, we can prove $D_L(\widetilde{P_1}, \widetilde{P_2}) = \infty$.

Proof of (h):

Compared to (g), we can notice that only the conditions for the controller 1 and 2 are flipped. Thus, by symmetry the proof is the same as (g).

Proof of (i):

Since $\Sigma_2 \geq 150$, we can find $k \geq 3$ such that

$$\frac{a^{2(k-2)} - 1}{1 - a^{-2}} \leq \frac{\Sigma_2}{24} < \frac{a^{2(k-1)} - 1}{1 - a^{-2}}. \tag{10.48}$$

Let $k_1 = 1$ and $k_2 = k$. By plugging these parameters into Lemma 5.3, the lower bound of Lemma 5.3 reduces to

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq (\sqrt{\frac{a^{2(k-1)}\frac{1-a^{-2(k-1)}}{1-a^{-2}}}{2^{2I'(\widetilde{P_1})}}} - \sqrt{\frac{a^{2(k-2)}(1 - a^{-(k-1)})^2}{(1 - a^{-1})^2}\widetilde{P_1}})_+^2 + 1. \tag{10.49}$$

Following the same steps as (10.45), we still have

$$I'(\widetilde{P_1}) \leq \frac{1}{2}\log e^{1.27656}. \tag{10.50}$$

Following the same steps as (10.27), we can prove

$$\frac{a^{2(k-2)}(1-a^{-(k-1)})^2}{(1-a^{-1})^2}\widetilde{P_1} \leq \frac{1}{5}\frac{a^{2(k-1)}(1-a^{-2(k-1)})}{(1-a^{-2})}. \tag{10.51}$$

Therefore, by plugging (10.50), (10.51) into (10.49), we conclude

$$\begin{aligned} D_L(\widetilde{P_1},\widetilde{P_2}) &\geq \frac{a^{2(k-1)}(1-a^{-2(k-1)})}{(1-a^{-2})}(\sqrt{\frac{1}{e^{1.27656}}}-\sqrt{\frac{1}{5}})^2+1 \\ &\geq \frac{a^{2(k-1)}(1-a^{-2(k-1)})}{(1-a^{-2})}0.00655882...+1 \\ &\overset{(A)}{\geq} \frac{\Sigma_2}{24}0.00655882...+1 \\ &\geq 0.000273284...\Sigma_2+1 \\ &\geq 0.0002732\Sigma_2+1. \end{aligned}$$

(A): This comes from (10.48).

Proof of (j):

We will prove this by analyzing the centralized controller performance which has both $y_1[n]$, $y_2[n]$ and has no input power constraint.

Define $y_1'[n] := x[n] + v_1'[n]$ and $y_2'[n] := x[n] + v_2'[n]$ where $v_1'[n] \sim \mathcal{N}(0,\sigma_1^2)$ and $v_2'[n] \sim \mathcal{N}(0,\sigma_1^2)$ are i.i.d. random variables. Since the costs of centralized controllers are monotone in the variances of observations, the cost of the centralized controller with the observations $y_1[n]$, $y_2[n]$ is larger than the cost of the centralized controller with the observations $y_1'[n]$, $y_2'[n]$. Moreover, by the maximum ratio combining, the cost of the centralized controller with the observations $y_1'[n]$, $y_2'[n]$ is equivalent to the cost of the centralized controller with a scalar observation $\frac{y_1'[n]+y_2'[n]}{2}$.

Now, we can apply Lemma 5.1 to analyze the performance of such a controller with the observation $\frac{y_1'[n]+y_2'[n]}{2}$. Let Σ_E be the Kalman filtering performance with the observation $\frac{y_1'[n]+y_2'[n]}{2}$. Then, by Lemma 5.1, Σ_E is lower bounded by

$$\begin{aligned} \Sigma_E &= \frac{(a^2-1)(\frac{\sigma_{v1}^2}{2})-1+\sqrt{((a^2-1)(\frac{\sigma_{v1}^2}{2})-1)^2+4a^2\frac{\sigma_{v1}^2}{2}}}{2a^2} \\ &\geq \frac{\max((a^2-1)\frac{\sigma_{v1}^2}{2}-1, \sqrt{2}a\sigma_{v1}-1)}{2a^2}. \end{aligned} \tag{10.52}$$

Therefore, for all $\widetilde{P_1}, \widetilde{P_2}$ the decentralized controller's cost is lower bounded as follows:

$$\begin{aligned}
D_L(\widetilde{P_1}, \widetilde{P_2}) &\overset{(A)}{\geq} \inf_{|a-k|<1} \frac{(2ak-k^2)\Sigma_E + 1}{1-(a-k)^2} \\
&= \inf_{|a-k|<1} \frac{2ak-k^2}{1-(a-k)^2}\Sigma_E + \frac{1}{1-(a-k)^2} \\
&\overset{(B)}{\geq} \inf_{|a-k|<1} \frac{1-a^2+2ak-k^2}{1-(a-k)^2}\Sigma_E + 1 \\
&= \Sigma_E + 1
\end{aligned} \tag{10.53}$$

(A): The decentralized control cost is larger than the centralized controller's cost with the observation $\frac{y_1'[n]+y_2'[n]}{2}$. Moreover, when $|a-k| \geq 1$ the centralized control system is unstable, and the cost diverges to infinity. When $|a-k| < 1$, the cost analysis follows from Lemma 5.1.
(B): This comes from $a > 1$ and $2ak - k^2 \geq 1 - (a-k)^2 > 0$.

Therefore, by (10.52) and (10.53) for all $\widetilde{P_1}, \widetilde{P_2}$ we have

$$\begin{aligned}
D_L(\widetilde{P_1}, \widetilde{P_2}) &\geq \max\left(\frac{\max((a^2-1)\frac{\sigma_{v1}^2}{2} - 1, \sqrt{2}a\sigma_{v1} - 1)}{2a^2}, 1\right) \\
&\geq \frac{\max((a^2-1)\frac{\sigma_{v1}^2}{2} - 1, \sqrt{2}a\sigma_{v1} - 1)}{4a^2} + \frac{1}{2} \\
&\geq \frac{\max((a^2-1)\frac{\sigma_{v1}^2}{2} - 1, \sqrt{2}a\sigma_{v1} - 1)}{4a^2} + \frac{1}{2a^2} \\
&\geq \frac{\max((a^2-1)\frac{\sigma_{v1}^2}{2}, \sqrt{2}a\sigma_{v1}, 1)}{4a^2}
\end{aligned}$$

By (10.9) we already know

$$\Sigma_1 \leq \frac{(1+\sqrt{2})\max(1, (a^2-1)\sigma_{v1}^2, 2a\sigma_{v1})}{2a^2}.$$

Therefore,

$$\begin{aligned}
D(\widetilde{P_1}, \widetilde{P_2}) &\geq \frac{\max(1, (a^2-1)\frac{\sigma_{v1}^2}{2}, \sqrt{2}a\sigma_{v1})}{4a^2} \\
&\geq \min\left(\frac{\frac{1}{4}}{\frac{1+\sqrt{2}}{2}}, \frac{\frac{1}{8}}{\frac{1+\sqrt{2}}{2}}, \frac{\frac{\sqrt{2}}{4}}{1+\sqrt{2}}\right)\Sigma_1 \\
&= \frac{1}{4(1+\sqrt{2})}\Sigma_1 \\
&\geq 0.1035\Sigma_1.
\end{aligned}$$

□

As mentioned in (10.53), $D(\widetilde{P_1}, \widetilde{P_2}) \geq 1$. Thus, the statement (j) is true.

Proof of Proposition 5.1 of page 249. Consider the power-distortion tradeoff $D(P_1, P_2)$ for the decentralized control problem shown in Problem K. Since we can achieve the tradeoff of the single controller systems by turning on only one controller, we have

$$(D(P_1, P_2), P_1, P_2) \leq (\min(D_{\sigma_{v1}}(P_1), D_{\sigma_{v2}}(P_2)), P_1, P_2)$$

where the definition of $D_\sigma(P)$ is shown in Problem L.

By Lemma 4.14 of Chapter 4, if there exists $c \geq 1$ such that for all $\widetilde{P_1}, \widetilde{P_2} \geq 0$,

$$\min(D_{\sigma 1}(c\widetilde{P_1}), D_{\sigma 2}(c\widetilde{P_2})) \leq c \cdot D_L(\widetilde{P_1}, \widetilde{P_2}),$$

then for all $q, r_1, r_2 \geq 0$ we have

$$\frac{\min_{P_1, P_2 \geq 0} q \min(D_{\sigma 1}(cP_1), D_{\sigma 2}(cP_2)) + r_1 P_1 + r_2 P_2}{\min_{\widetilde{P_1}, \widetilde{P_2} \geq 0} q D_L(\widetilde{P_1}, \widetilde{P_2}) + r_1 P_1 + r_2 P_2} \leq c$$

which finishes the proof. Therefore, we will only prove that such c exists.

Before we start the proof, define the subscript max as $argmax_{i \in \{1,2\}} \widetilde{P_i}$. For example, if $\widetilde{P_1} < \widetilde{P_2}$ then $\widetilde{P_{max}} = \widetilde{P_2}$, $P_{max} = P_2$, $\Sigma_{max} = \Sigma_2$, $D_{\sigma_{vmax}}(P) = D_{\sigma_{v2}}(P)$ and so on. Furthermore, for notational simplicity, we write $D_{\sigma_{v1}}(\cdot), D_{\sigma_{v2}}(\cdot), D_{\sigma_{vmax}}(\cdot)$ as $D_{v1}(\cdot), D_{v2}(\cdot), D_{vmax}(\cdot)$ respectively.

For the proof, we will first divide the cases based on Σ_1, Σ_2 then further divide based on $\widetilde{P_1}, \widetilde{P_2}$. Remind that since $\sigma_{v1} \leq \sigma_{v2}$, we have $\Sigma_1 \leq \Sigma_2$. We can use this fact to reduce the cases.

(i) When $\Sigma_1 \leq \Sigma_2 \leq 150$

(i-i) When $\frac{1}{150} \leq \max(\widetilde{P_1}, \widetilde{P_2})$

Lower bound: By Corollary 5.4 (j)

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq 1$$

Upper bound:

If $(a^2 - 1) \leq \frac{1}{\max(1, 7.25\Sigma_{max})}$, then the range for t in Corollary 5.2 (ii') is not an empty set. Therefore, by plugging $t = \frac{8}{\max(1, 7.25\Sigma_E)}$ we get

$$\begin{aligned}(D_{\sigma max}(P_{max}), P_{max}) &\leq (\frac{49}{8}\max(2, 14.5\Sigma_{max}), \frac{8}{\max(1, 7.25\Sigma_{max})}) \\ &\leq (\frac{49}{8} \cdot 14.5 \cdot 150, 8)(\because \Sigma_1 \leq \Sigma_2 \leq 150).\end{aligned}$$

If $(a^2 - 1) \geq \frac{1}{\max(1, 7.25\Sigma_{max})}$, by Corollary 5.2 (i') we get

$$\begin{aligned}(D_{\sigma max}(P_{max}), P_{max}) &\leq (7.25\Sigma_{max} + \frac{6.25}{a^2 - 1}, (a^2 - 1)^2\Sigma_{max} + (a^2 - 1)) \\ &\leq (7.25\Sigma_{max} + \frac{6.25}{a^2 - 1}, 27.5625\Sigma_{max} + 5.25)(\because 1 < |a| \leq 2.5) \\ &\leq (7.25\Sigma_{max} + 6.25\max(1, 7.25\Sigma_{max}), 27.5625\Sigma_{max} + 5.25) \\ &\leq (7.25 \cdot 150 + 6.25 \cdot 7.25 \cdot 150, 27.5625 \cdot 150 + 5.25)(\because \Sigma_1 \leq \Sigma_2 \leq 150) \\ &\leq (7884.375, 4139.625).\end{aligned}$$

Ratio: c is upper bounded by

$$c \leq \frac{4139.625}{\frac{1}{150}} < 10^6.$$

(i-ii) When $\frac{1}{20}(a^2-1) \leq \max(\widetilde{P_1}, \widetilde{P_2}) \leq \frac{1}{150}$

Lower bound: By Corollary 5.4 (d),

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq 0.00389 \frac{1}{\max(\widetilde{P_1}, \widetilde{P_2})} + 1. \tag{10.54}$$

Upper bound:

If $8(a^2-1) \leq \max(\widetilde{P_1}, \widetilde{P_2}) \leq \frac{8}{\max(1, 7.25\Sigma_{max})}$, then we can put $t = \widetilde{P_{max}}$ in Corollary 5.2 (ii') for $D_{\sigma max}(P_{max})$. Therefore, we get

$$(D_{\sigma max}(P_{max}), P_{max}) \leq (\frac{49}{\widetilde{P_{max}}}, \widetilde{P_{max}}).$$

If $\frac{1}{20}(a^2-1) \leq \max(\widetilde{P_1}, \widetilde{P_2}) \leq 8(a^2-1)$

In this case, the lower bound (10.54) can be further lower bounded as

$$(D_L(P_1, P_2), P_{max}) \geq (\frac{0.00389}{24.5(a^2-1)} + 1, \frac{1}{20}(a^2-1)).$$

By Corollary 5.2 (i'), we have

$$\begin{aligned}
(D_{\sigma max}(P_{max}), P_{max}) &\leq (7.25\Sigma_{max} + \frac{6.25}{a^2-1}, (a^2-1)^2\Sigma_{max} + (a^2-1)) \\
&\leq (7.25 \cdot 150 + \frac{6.25}{a^2-1}, 5.25 \cdot 150(a^2-1) + (a^2-1))(\because 1 \leq |a| < 2.5, \Sigma_1 \leq \Sigma_2 \leq 150) \\
&= (\frac{6.25}{a^2-1} + 1087.5, 788.5(a^2-1)).
\end{aligned}$$

If $\frac{8}{\max(1, 7.25\Sigma_{max})} \leq \max(\widetilde{P_1}, \widetilde{P_2}) \leq \frac{1}{150}$

Notice that this case never happens since

$$\frac{8}{\max(1, 7.25\Sigma_{max})} \geq \frac{8}{7.25 \cdot 150} > \frac{1}{150}.$$

Ratio: c is upper bounded by

$$c \leq \frac{24.5 \times 6.25}{0.00389} < 40000.$$

(i-iii) When $\max(\widetilde{P_1}, \widetilde{P_2}) \leq \frac{1}{20}(a^2-1)$

Lower bound: By Corollary 5.4 (c),

$$D_L(\widetilde{P_1}, \widetilde{P_2}) = \infty.$$

We do not need a corresponding upper bound.

(ii) When $\Sigma_1 \leq 150 \leq \Sigma_2$

(ii-i) When $\frac{20}{a^2-1} \geq \Sigma_2$

(ii-i-i) When $\frac{1}{150} \leq \widetilde{P_1}$

Lower bound: By Corollary 5.4 (j),

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq 1$$

If $(a^2-1) \leq \frac{1}{\max(1,7.25\Sigma_1)}$

Upper bound: By putting $t = \frac{8}{\max(1,7.25\Sigma_1)}$ to Corollary 5.2 (ii') for $D_{\sigma 1}(P_1)$, we have

$$\begin{aligned}(D_{\sigma 1}(P_1), P_1) &\leq (\frac{49}{8}\max(1, 7.25\Sigma_1), \frac{8}{\max(1, 7.25\Sigma_1)}) \\ &\leq (\frac{49}{8} \cdot 7.25 \cdot 150, 8)(\because \Sigma_1 \leq 150)\end{aligned}$$

If $(a^2-1) \geq \frac{1}{\max(1,7.25\Sigma_1)}$

Upper bound: By Corollary 5.2 (i'),

$$\begin{aligned}(D_{\sigma 1}(P_1), P_1) &\leq (7.25\Sigma_1 + \frac{6.25}{a^2-1}, 27.5625\Sigma_1 + 5.25) \\ &\leq (7.25 \cdot 150 + 6.25\max(1, 7.25\Sigma_1), 27.5625 \cdot 150 + 5.25)(\because \Sigma_1 \leq 150) \\ &\leq (7.25 \cdot 150 + 6.25 \cdot 7.25 \cdot 150, 27.5625 \cdot 150 + 5.25).\end{aligned}$$

Ratio: c is upper bounded by

$$c \leq \frac{27.5625 \cdot 150 + 5.25}{\frac{1}{150}} = 620943.75.$$

(ii-i-ii) When $\frac{1}{\Sigma_2} \leq \widetilde{P_1} \leq \frac{1}{150}$

Lower bound: By Corollary 5.4 (f)

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq \frac{0.0006976}{\widetilde{P_1}} + 1. \tag{10.55}$$

If $\frac{8}{\max(1,7.25\Sigma_1)} \leq \widetilde{P_1} \leq \frac{1}{150}$,

This never happens since $\frac{8}{\max(1,7.25\Sigma_1)} \geq \frac{8}{7.25 \cdot 150} > \frac{1}{150}$.

If $8(a^2-1) \leq \widetilde{P_1} \leq \frac{8}{\max(1,7.25\Sigma_1)}$,

Upper bound: By plugging $t = \widetilde{P_1}$ to Corollary 5.2 (ii') for $D_{\sigma 1}(P_1)$, we get

$$(D_{\sigma 1}(P_1), P_1) \leq (\frac{49}{\widetilde{P_1}}, \widetilde{P_1}).$$

If $\frac{1}{\Sigma_2} \leq \widetilde{P_1} \leq 8(a^2-1)$,

Here the lower bound of (10.55) is further lower bounded by

$$(D_L(\widetilde{P_1}, \widetilde{P_2}), \widetilde{P_1}) \geq (\frac{0.0006976}{8(a^2-1)} + 1, \frac{1}{\Sigma_2}).(\because \frac{1}{\Sigma_2} \leq \widetilde{P_1} \leq 8(a^2-1)).$$

Upper bound: When $\frac{1}{\Sigma_2} \leq \widetilde{P_1} \leq 8(a^2-1)$ and $(a^2-1) \leq \frac{1}{\max(1,7.25\Sigma_1)}$, we can plug $t = 8(a^2-1)$ for $D_{\sigma 1}(P_1)$ to Corollary 5.2 (ii') for $D_{\sigma 1}(P_1)$. Then, we get

$$\begin{aligned}(D_{\sigma 1}(P_1), P_1) &\leq (\frac{49}{8(a^2-1)}, 8(a^2-1)) \\ &\leq (\frac{49}{8(a^2-1)}, \frac{8\cdot 20}{\Sigma_2})(\because \text{In (ii-i), we assumed } \frac{20}{a^2-1} \geq \Sigma_2).\end{aligned}$$

When $\frac{1}{\Sigma_2} \leq P_1 \leq 8(a^2-1)$ and $(a^2-1) > \frac{1}{\max(1,7.25\Sigma_1)}$, by Corollary 5.2 (i') we get

$$\begin{aligned}(D_{\sigma 1}(P_1), P_1) &\leq (7.25\Sigma_1 + \frac{6.25}{a^2-1}, (a^2-1)^2\Sigma_1 + (a^2-1)) \\ &\leq (7.25\Sigma_1 + \frac{6.25}{a^2-1}, \frac{20^2\Sigma_1}{\Sigma_2^2} + \frac{20}{\Sigma_2}) \\ &(\because \text{In (ii-i), we assumed } \frac{20}{a^2-1} \geq \Sigma_2) \\ &\leq (\frac{6.25}{a^2-1} + 7.25\cdot 150, \frac{20^2+20}{\Sigma_2}).(\because \Sigma_1 \leq 150 \leq \Sigma_2)\end{aligned}$$

Ratio: c is upper bounded by

$$c \leq \frac{8\times 6.25}{0.0006976} < 72000.$$

(ii-i-iii) When $\widetilde{P_1} \leq \frac{1}{\Sigma_2}$ and $\max(\widetilde{P_1}, \widetilde{P_2}) = \widetilde{P_2} > \frac{1}{\Sigma_2}$

Lower bound: By Corollary 5.4 (e)

$$D_L(P_1, P_2) \geq 0.0006976\Sigma_2 + 1.$$

First, since $\Sigma_2 \geq 150$, we can see that $\max(1, 7.25\Sigma_2) = 7.25\Sigma_2$.

If $(a^2-1) \leq \frac{1}{7.25\Sigma 2}$

Upper bound: By plugging $t = \frac{8}{7.25\Sigma_2}$ into Corollary 5.2 (ii') for $D_{\sigma 2}(P_2)$, we get

$$(D_{\sigma 2}(P_2), P_2) \leq (\frac{49}{8}\cdot 7.25\Sigma_2, \frac{8}{7.25\Sigma_2}).$$

If $(a^2-1) \geq \frac{1}{7.25\Sigma 2}$

Upper bound: By Corollary 5.2 (i'), we get

$$\begin{aligned}(D_{\sigma 2}(P_2), P_2) &\leq (7.25\Sigma_2 + \frac{6.25}{a^2-1}, (a^2-1)^2\Sigma_2 + (a^2-1)) \\ &\leq (7.25\Sigma_2 + \frac{6.25}{a^2-1}, \frac{20^2\Sigma_2}{\Sigma_2^2} + \frac{20}{\Sigma_2}) \\ &(\because \text{In (ii-i), we assumed } \frac{20}{a^2-1} \geq \Sigma_2) \\ &\leq (7.25\Sigma_2 + 6.25\cdot 7.25\Sigma_2, \frac{20^2+20}{\Sigma_2}).\end{aligned}$$

Ratio: c is upper bounded by

$$c \leq \frac{7.25 + 6.25\times 7.25}{0.0006976} < 76000.$$

(ii-i-iv) When $\widetilde{P_1} \leq \frac{1}{\Sigma_2}$ and $\frac{1}{20}(a^2-1) \leq \max(\widetilde{P_1}, \widetilde{P_2}) \leq \frac{1}{\Sigma_2}$

Lower bound: By Corollary 5.4 (d), we have

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq \frac{0.00389}{\max(\widetilde{P_1}, \widetilde{P_2})} + 1. \tag{10.56}$$

If $\frac{8}{\max(1, 7.25\Sigma_{max})} \leq \max(\widetilde{P_1}, \widetilde{P_2}) \leq \frac{1}{\Sigma_2}$

This case never happens, since $\frac{8}{\max(1,7.25\Sigma_{max})} = \frac{8}{7.25\Sigma_2} > \frac{1}{\Sigma_2}$.

If $8(a^2-1) \leq \max(\widetilde{P_1}, \widetilde{P_2}) \leq \frac{8}{\max(1,7.25\Sigma_{max})}$

Upper bound: By plugging $t = \widetilde{P_{max}}$ into Corollary 5.2 (ii') for $D_{\sigma max}(P_{max})$, we have

$$(D_{\sigma max}(P_{max}), P_{max}) \leq (\frac{49}{\widetilde{P_{max}}}, \widetilde{P_{max}}).$$

If $\frac{1}{20}(a^2-1) \leq \max(P_1, P_2) \leq 8(a^2-1)$

In this case, the lower bound of (10.56) is further lower bounded by

$$(D_L(\widetilde{P_1}, \widetilde{P_2}), P_{max}) \geq (\frac{0.00389}{8(a^2-1)} + 1, \frac{1}{20}(a^2-1)).$$

Upper bound: By Corollary 5.2 (i'), we have

$$\begin{aligned}
(D_{\sigma max}(P_{max}), P_{max}) &\leq (7.25\Sigma_{max} + \frac{6.25}{a^2-1}, (a^2-1)^2\Sigma_{max} + (a^2-1)) \\
&\leq (7.25\Sigma_2 + \frac{6.25}{a^2-1}, (a^2-1)^2\Sigma_2 + (a^2-1))(\because \Sigma_1 \leq \Sigma_2) \\
&\leq (\frac{7.25 \cdot 20}{a^2-1} + \frac{6.25}{a^2-1}, 20(a^2-1) + (a^2-1))(\because \text{In (ii-i), we assumed } \frac{20}{a^2-1} \geq \Sigma_2) \\
&\leq (\frac{7.25 \cdot 20 + 6.25}{a^2-1}, 21(a^2-1))
\end{aligned}$$

Ratio: c is upper bounded by

$$c \leq \frac{7.25 \cdot 20 + 6.25}{\frac{0.00389}{8}} < 320000.$$

(ii-i-v) When $\widetilde{P_1} \leq \frac{1}{\Sigma_2}$ and $\max(\widetilde{P_1}, \widetilde{P_2}) \leq \frac{1}{20}(a^2-1)$

Lower bound: By Corollary 5.4 (c),

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq \infty.$$

Thus, we do not need a corresponding upper bound in this case.

(ii-ii) When $\frac{20}{a^2-1} \leq \Sigma_2$

(ii-ii-i) When $\frac{1}{150} \leq \widetilde{P_1}$

Compared to (ii-i-i), the only difference is Σ_2 and Σ_2 does not affect the result of (ii-i-i). Therefore, in the same way as (ii-i-i), we can prove that c is bounded by the same constant as (ii-i-i).

(ii-ii-ii) When $\frac{1}{20}(a^2-1) \leq \widetilde{P_1} \leq \frac{1}{150}$

Lower bound: Since in (ii-ii) we assumed $\frac{20}{a^2-1} \leq \Sigma_2$, we have $\frac{1}{\Sigma_2} \leq \frac{1}{20}(a^2-1) \leq P_1 \leq \frac{1}{150}$. Therefore, we can apply Corollary 5.4 (f) to get

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq \frac{0.0006976}{\widetilde{P_1}} + 1. \tag{10.57}$$

If $\frac{8}{\max(1,7.25\Sigma_1)} \leq \widetilde{P_1} \leq \frac{1}{150}$

Since we assumed $\Sigma_1 \leq 150$ in (ii), $\frac{8}{\max(1,7.25\Sigma_1)} \geq \frac{8}{7.25 \cdot 150} > \frac{1}{150}$. Therefore, this case never happens.

If $8(a^2-1) \leq \widetilde{P_1} \leq \frac{8}{\max(1,7.25\Sigma_1)}$

Upper bound: By plugging $t = \widetilde{P_1}$ into Corollary 5.2 (ii') for $D_{\sigma 1}(P_1)$, we have

$$(D_{\sigma 1}(P_1), P_1) \leq (\frac{49}{\widetilde{P_1}}, \widetilde{P_1}).$$

If $\frac{1}{20}(a^2-1) \leq \widetilde{P_1} \leq 8(a^2-1)$

In this case, the lower bound of (10.57) is further lower bounded by

$$(D_L(\widetilde{P_1}, \widetilde{P_2}), \widetilde{P_1}) \geq (\frac{0.0006976}{8(a^2-1)} + 1, \frac{1}{20}(a^2-1)).$$

Upper bound: By Corollary 5.2 (i')

$$\begin{aligned}
(D_U(P_1), P_1) &\leq (7.25\Sigma_1 + \frac{6.25}{a^2-1}, (a^2-1)^2\Sigma_1 + (a^2-1)) \\
&\leq (7.25 \cdot 150 + \frac{6.25}{a^2-1}, 5.25 \cdot 150 \cdot (a^2-1) + (a^2-1)) \\
&(\because \Sigma_1 \leq 150, 1 < |a| \leq 2.5)
\end{aligned}$$

Ratio: c is upper bounded by

$$c \leq \frac{6.25 \times 8}{0.0006976} < 72000.$$

(ii-ii-iii) When $\widetilde{P_1} \leq \frac{1}{20}(a^2-1)$ and $\widetilde{P_2} \geq \frac{(a^2-1)^2\Sigma_2}{40000}$

Lower bound: By Corollary 5.4 (i), we have

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq 0.0002732\Sigma_2 + 1.$$

Upper bound: By Corollary 5.2 (i'), we have

$$\begin{aligned}
(D_{\sigma 2}(P_2), P_2) &\leq (7.25\Sigma_2 + \frac{6.25}{a^2-1}, (a^2-1)^2\Sigma_2 + (a^2-1)) \\
&\leq (7.25\Sigma_2 + \frac{6.25}{20}\Sigma_2, (a^2-1)^2\Sigma_2 + (a^2-1)^2\frac{\Sigma_2}{20}) \\
&(\because \text{In (ii-ii), we assumed } \frac{20}{a^2-1} \leq \Sigma_2)
\end{aligned}$$

Ratio: c is upper bounded by

$$c \leq \frac{1 + \frac{1}{20}}{\frac{1}{40000}} \leq 42000.$$

(ii-ii-iv) When $\widetilde{P_1} \le \frac{1}{20}(a^2-1)$ and $\widetilde{P_2} \le \frac{(a^2-1)^2\Sigma_2}{40000}$

Lower bound: By Corollary 5.2 (g), we have

$$D_L(\widetilde{P_1}, \widetilde{P_2}) = \infty$$

We do not need a matching upper bound.

(iii) When $150 \le \Sigma_1 \le \Sigma_2$

In this case, we can see that $\max(1, 7.25\Sigma_1) = 7.25\Sigma_1$, $\max(1, 7.25\Sigma_2) = 7.25\Sigma_2$.

(iii-i) When $\frac{20}{a^2-1} \ge \Sigma_1$ and $\frac{20}{a^2-1} \ge \Sigma_2$

(iii-i-i) When $\frac{1}{\Sigma_1} \le \widetilde{P_1}$

Lower bound: By Corollary 5.4 (j), we have

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \ge 0.1035\Sigma_1.$$

If $(a^2-1) \le \frac{1}{7.25\Sigma_1}$

Upper bound: By plugging $t = \frac{8}{7.25\Sigma_1}$ into Corollary 5.2 (ii') for $D_{\sigma 1}(P_1)$, we get

$$(D_{\sigma 1}(P_1), P_1) = (49 \cdot \frac{7.25\Sigma_1}{8}, \frac{8}{7.25\Sigma_1}).$$

If $(a^2-1) \ge \frac{1}{7.25\Sigma_1}$

Upper bound: By Corollary 5.2 (i'), we have

$$\begin{aligned}(D_{\sigma 1}(P_1), P_1) &= (7.25\Sigma_1 + \frac{6.25}{a^2-1}, (a^2-1)^2\Sigma_1 + (a^2-1)) \\ &\le (7.25\Sigma_1 + 6.25 \cdot 7.25\Sigma_1, (\frac{20}{\Sigma_1})^2\Sigma_1 + \frac{20}{\Sigma_1}) \\ &(\because \text{In (iii-i), we assumed } \frac{20}{a^2-1} \ge \Sigma_1) \\ &= (7.25^2\Sigma_1, \frac{20 \cdot 21}{\Sigma_1}).\end{aligned}$$

Ratio: c is upper bounded by

$$c \le \frac{7.25^2}{0.1035} < 510.$$

(iii-i-ii) When $\frac{1}{\Sigma_2} \le \widetilde{P_1} \le \frac{1}{\Sigma_1}$

Lower bound: Since in (iii) we assumed $150 \le \Sigma_1$, we have $\frac{1}{\Sigma_2} \le \widetilde{P_1} \le \frac{1}{\Sigma_1} \le \frac{1}{150}$. Therefore, we can apply Corollary 5.4 (f) to conclude

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \ge \frac{0.0006976}{\widetilde{P_1}} + 1. \tag{10.58}$$

If $\frac{8}{7.25\Sigma_1} \le \widetilde{P_1} \le \frac{1}{\Sigma_1}$

Since $\frac{8}{7.25\Sigma_1} > \frac{1}{\Sigma_1}$, this case never happens.

If $8(a^2-1) \le \widetilde{P_1} \le \frac{8}{7.25\Sigma_1}$

Upper bound: By plugging $t = \widetilde{P_1}$ into Corollary 5.2 (ii') for $D_{\sigma 1}(P_1)$, we have

$$(D_{\sigma 1}(P_1), P_1) \leq (\frac{49}{\widetilde{P_1}}, \widetilde{P_1}).$$

If $\frac{1}{\Sigma_2} \leq \widetilde{P_1} \leq 8(a^2-1)$

In this case, the lower bound of (10.58) is further lower bounded by

$$(D_L(\widetilde{P_1}, \widetilde{P_2}), \widetilde{P_1}) \geq (\frac{0.0006976}{8(a^2-1)} + 1, \frac{1}{\Sigma_2}).$$

If $\frac{1}{\Sigma_2} \leq P_1 \leq 24.5(a^2-1)$ and $(a^2-1) \leq \frac{1}{7.25\Sigma_1}$

Upper bound: By plugging $t = 8(a^2-1)$ into Corollary 5.2 (ii') for $D_{\sigma 1}(P_1)$, we have

$$\begin{aligned}
(D_{\sigma 1}(P_1), P_1) &\leq (\frac{49}{8(a^2-1)}, 8(a^2-1)) \\
&\leq (\frac{49}{8(a^2-1)}, \frac{8 \cdot 20}{\Sigma_2}) \\
&(\because \text{In (iii-i), we assumed } \frac{20}{a^2-1} \geq \Sigma_2)
\end{aligned}$$

If $\frac{1}{\Sigma_2} \leq P_1 \leq 8(a^2-1)$ and $(a^2-1) > \frac{1}{7.25\Sigma_1}$

Upper bound: By Corollary 5.2 (i'), we have

$$\begin{aligned}
(D_{\sigma 1}(P_1), P_1) &\leq (7.25\Sigma_1 + \frac{6.25}{a^2-1}, (a^2-1)^2\Sigma_1 + (a^2-1)) \\
&\leq (\frac{7.25 \cdot 20}{a^2-1} + \frac{6.25}{a^2-1}, \frac{20^2\Sigma_1}{\Sigma_2^2} + \frac{20}{\Sigma_2}) \\
&(\because \text{In (iii-i), we assumed } \Sigma_1 \leq \frac{20}{a^2-1}, \Sigma_2 \leq \frac{20}{a^2-1}) \\
&\leq (\frac{7.25 \cdot 20 + 6.25}{a^2-1}, \frac{20^2+20}{\Sigma_2}) \\
&(\because \Sigma_1 \leq \Sigma_2)
\end{aligned}$$

Ratio: c is upper bounded by

$$c \leq \frac{7.25 \cdot 20 + 6.25}{\frac{0.0006976}{8}} \leq 2 \times 10^6.$$

(iii-i-iii) When $\widetilde{P_1} \leq \frac{1}{\Sigma_2}$ and $\max(\widetilde{P_1}, \widetilde{P_2}) = \widetilde{P_2} > \frac{1}{\Sigma_2}$

Lower bound: By Corollary 5.4 (e),

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq 0.0006976\Sigma_2 + 1.$$

If $(a^2-1) \leq \frac{1}{7.25\Sigma_2}$

Upper bound: By plugging $t = \frac{8}{7.25\Sigma_2}$ into Corollary 5.2 (ii') for $D_{\sigma 2}(P_2)$, we get

$$(D_{\sigma 2}(P_2), P_2) \leq (14.5\Sigma_2, \frac{24.5}{7.25\Sigma_2}).$$

If $(a^2-1) \geq \frac{1}{7.25\Sigma_2}$

Upper bound: By Corollary 5.2 (i'), we have

$$\begin{aligned}(D_{\sigma 2}(P_2), P_2) &\leq (7.25\Sigma_2 + \frac{6.25}{a^2-1}, (a^2-1)^2\Sigma_2 + (a^2-1)) \\ &\leq (7.25\Sigma_2 + 6.25 \cdot 7.25\Sigma_2, \frac{20^2}{\Sigma_2} + \frac{20}{\Sigma_2}) \\ &(\because \text{In (iii-i), we assumed } \Sigma_2 \leq \frac{20}{a^2-1}) \\ &\leq (7.25^2\Sigma_2, \frac{20 \cdot 21}{\Sigma_2}).\end{aligned}$$

Ratio: c is upper bounded by

$$c \leq \frac{7.25^2}{0.0006976} < 80000.$$

(iii-i-iv) When $\widetilde{P_1} \leq \frac{1}{\Sigma_2}$ and $\frac{1}{20}(a^2-1) \leq \max(\widetilde{P_1}, \widetilde{P_2}) \leq \frac{1}{\Sigma_2}$

Lower bound: Since we assumed $\Sigma_2 \geq 150$ in (iii), we have $\max(P_1, P_2) \leq \frac{1}{\Sigma_2} \leq \frac{1}{150} \leq \frac{1}{75}$. Therefore, by Corollary 5.4 (d) we can see

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq \frac{0.00389}{\max(\widetilde{P_1}, \widetilde{P_2})} + 1. \tag{10.59}$$

If $\frac{8}{7.25\Sigma_{max}} \leq \max(\widetilde{P_1}, \widetilde{P_2}) \leq \frac{1}{\Sigma_2}$

This never happens since

$$\frac{8}{7.25\Sigma_{max}} > \frac{1}{\Sigma_{max}} \geq \frac{1}{\Sigma_2}.$$

If $8(a^2-1) \leq \max(\widetilde{P_1}, \widetilde{P_2}) \leq \frac{8}{7.25\Sigma_{max}}$

Upper bound: By plugging $t = \widetilde{P_{max}}$ into Corollary 5.2 (ii'), we get

$$(D_{\sigma max}(P_{max}), P_{max}) \leq (\frac{49}{\widetilde{P_{max}}}, \widetilde{P_{max}}).$$

If $\frac{1}{20}(a^2-1) \leq \max(\widetilde{P_1}, \widetilde{P_2}) \leq 8(a^2-1)$

In this case, we can notice that the lower bound of (10.59) is further lower bounded by

$$(D_L(\widetilde{P_1}, \widetilde{P_2}), \widetilde{P_{max}}) \geq (\frac{0.00389}{8(a^2-1)} + 1, \frac{1}{20}(a^2-1))$$

Upper bound: By Corollary 5.2 (i'), we get

$$\begin{aligned}(D_{\sigma max}(P_{max}), P_{max}) &\leq (7.25\Sigma_{max} + \frac{6.25}{a^2-1}, (a^2-1)^2\Sigma_{max} + (a^2-1)) \\ &\leq (\frac{7.25 \cdot 20}{a^2-1} + \frac{6.25}{a^2-1}, 20(a^2-1) + (a^2-1)) \\ &(\because \text{In (iii-i), we assumed } \Sigma_1 \leq \frac{20}{a^2-1}, \Sigma_2 \leq \frac{20}{a^2-1}) \\ &\leq (\frac{7.25 \cdot 20 + 6.25}{a^2-1}, 21(a^2-1))\end{aligned}$$

Ratio: c is upper bounded by

$$c \leq \frac{7.25 \cdot 20 + 6.25}{\frac{0.00389}{8}} < 320000.$$

(iii-i-v) When $\widetilde{P_1} \leq \frac{1}{\Sigma_2}$ and $\max(\widetilde{P_1}, \widetilde{P_2}) \leq \frac{1}{20}(a^2-1)$

Lower bound: By Corollary 5.4 (c),

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq \infty.$$

We do not need a corresponding upper bound.

(iii-ii) When $\Sigma_1 \leq \frac{20}{a^2-1} \leq \Sigma_2$

(iii-ii-i) When $\frac{1}{\Sigma_1} \leq \widetilde{P_1}$

Compared to the case (iii-i-i), the conditions for Σ_1, $\widetilde{P_1}$ are the same and the only difference is the condition for Σ_2. However, the condition for Σ_2 does not affect the argument of (iii-i-i). Thus, the same bound on c as (iii-i-i) still holds for this case.

(iii-ii-ii) When $\frac{1}{20}(a^2-1) \leq \widetilde{P_1} \leq \frac{1}{\Sigma_1}$

Lower bound: Since we assumed $150 \leq \Sigma_1$ in (iii), we have $\frac{1}{\Sigma_2} \leq \frac{1}{20}(a^2-1) \leq P_1 \leq \frac{1}{\Sigma_1} \leq \frac{1}{150}$. Thus, we can apply Corollary 5.4 (f) to get

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq \frac{0.0006976}{\widetilde{P_1}} + 1. \tag{10.60}$$

If $\frac{8}{7.25\Sigma_1} \leq \widetilde{P_1} \leq \frac{1}{\Sigma_1}$

This case never happens.

If $8(a^2-1) \leq \widetilde{P_1} \leq \frac{8}{7.25\Sigma_1}$

Upper bound: By plugging $t = \widetilde{P_1}$ into Corollary 5.2 (ii') for $D_{\sigma 1}(P_1)$, we get

$$(D_{\sigma 1}(P_1), P_1) \leq (\frac{49}{\widetilde{P_1}}, \widetilde{P_1}).$$

If $\frac{1}{20}(a^2-1) \leq P_1 \leq 8(a^2-1)$

In this case, the lower bound of (10.60) can be further lower bounded as

$$(D_L(\widetilde{P_1}, \widetilde{P_2}), \widetilde{P_1}) \geq (\frac{0.0006976}{8(a^2-1)} + 1, \frac{1}{20}(a^2-1))$$

Upper bound: By Corollary 5.2 (i'), we have

$$\begin{aligned}
(D_{\sigma 1}(P_1), P_1) &\leq (7.25\Sigma_1 + \frac{6.25}{a^2-1}, (a^2-1)^2\Sigma_1 + (a^2-1)) \\
&\leq (\frac{7.25 \cdot 20}{a^2-1} + \frac{6.25}{a^2-1}, 20(a^2-1) + (a^2-1)) \\
&\quad (\because \text{In (iii-ii), we assumed } \Sigma_1 \leq \frac{20}{a^2-1}.) \\
&\leq (\frac{7.25 \cdot 20 + 6.25}{a^2-1}, 21(a^2-1))
\end{aligned}$$

Ratio: c is upper bounded by

$$c \leq \frac{7.25 \cdot 20 + 6.25}{\frac{0.006976}{8}} < 320000.$$

(iii-ii-iii) When $\widetilde{P_1} \leq \frac{1}{20}(a^2 - 1)$ and $\widetilde{P_2} \geq \frac{(a^2-1)^2\Sigma_2}{40000}$
Lower bound: By Corollary 5.4 (i), we have

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq 0.0002732\Sigma_2 + 1.$$

Upper bound: By Corollary 5.2 (i'), we have

$$\begin{aligned}(D_{\sigma 2}(P_2), P_2) &\leq (7.25\Sigma_2 + \frac{6.25}{a^2-1}, (a^2-1)^2\Sigma_2 + (a^2-1)) \\ &\leq (7.25\Sigma_2 + \frac{6.25}{20}\Sigma_2, (a^2-1)^2\Sigma_2 + (a^2-1)^2\frac{\Sigma_2}{20}) \\ &(\because \text{In (iii-ii), we assumed } \frac{20}{a^2-1} \leq \Sigma_2)\end{aligned}$$

Ratio: c is upper bounded by

$$c \leq 40000(1 + \frac{1}{20}) \leq 42000.$$

(iii-ii-iv) When $\widetilde{P_1} \leq \frac{1}{20}(a^2 - 1)$ and $\widetilde{P_2} \leq \frac{(a^2-1)^2\Sigma_2}{40000}$
Lower bound: By Corollary 5.4 (g),

$$D_L(\widetilde{P_1}, \widetilde{P_2}) = \infty.$$

Therefore, we do not need a corresponding upper bound.
(iii-iii) When $\frac{20}{a^2-1} \leq \Sigma_1 \leq \Sigma_2$
(iii-iii-i) When $\widetilde{P_1} \geq \frac{(a^2-1)^2\Sigma_1}{40000}$
Lower bound: By Corollary 5.4 (j), we have

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq 0.1035\Sigma_1$$

Upper bound: By Corollary 5.2 (i'), we have

$$\begin{aligned}(D_{\sigma 1}(P_1), P_1) &\leq (7.25\Sigma_1 + \frac{6.25}{a^2-1}, (a^2-1)^2\Sigma_1 + (a^2-1)) \\ &\leq (7.25\Sigma_1 + \frac{6.25}{20}\Sigma_1, (a^2-1)^2\Sigma_1 + (a^2-1)^2\frac{\Sigma_1}{20}) \\ &(\because \text{In (iii-iii), we assumed } \frac{20}{a^2-1} \leq \Sigma_1)\end{aligned}$$

Ratio: c is upper bounded by

$$c \leq 40000(1 + \frac{1}{20}) \leq 42000.$$

(iii-iii-ii) When $\widetilde{P_1} \leq \frac{(a^2-1)^2\Sigma_1}{40000}$ and $\widetilde{P_2} \geq \frac{(a^2-1)^2\Sigma_2}{40000}$

Lower bound: By Corollary 5.4 (b), we have

$$D_L(P_1, P_2) \geq 0.002774\Sigma_2 + 1.$$

Upper bound: By Corollary 5.2 (i'), we have

$$\begin{aligned}(D_U(P_2), P_2) &\leq (7.25\Sigma_2 + \frac{6.25}{a^2-1}, (a^2-1)^2\Sigma_2 + (a^2-1)) \\ &\leq (7.25\Sigma_2 + \frac{6.25}{20}\Sigma_2, (a^2-1)^2\Sigma_2 + (a^2-1)^2\frac{\Sigma_2}{20}) \\ &(\because \text{In (iii-iii), we assume } \frac{20}{a^2-1} \leq \Sigma_2)\end{aligned}$$

Ratio: c is upper bounded by

$$c \leq 40000(1 + \frac{1}{20}) \leq 42000.$$

(iii-iii-iii) When $\widetilde{P_1} \leq \frac{(a^2-1)^2\Sigma_1}{40000}$ and $\widetilde{P_2} \leq \frac{(a^2-1)^2\Sigma_2}{40000}$
Lower bound: By Corollary 5.4 (a), we have

$$D_L(\widetilde{P_1}, \widetilde{P_2}) = \infty.$$

Therefore, we do not need a corresponding upper bound.
Finally, by (i), (ii), (iii), we get the constant $c \leq 6 \times 10^6$ and prove the proposition. □

10.3 Proof of Lemma 5.4, Corollary 5.1 and Proposition 5.2

Proof of Lemma 5.4 of Page 249. For simplicity, we assume $a = 1$, $1 < k_1 < k_2 < k$. The remaining cases when $a = -1$ or $k_1 = 1$ or $k_2 = k_1$ or $k = k_2$ easily follow with minor modifications.

We essentially follow the proof of Lemma 5.3. However, since $|a| = 1$, the sum of the sequence, $\frac{1}{|a|}, \frac{1}{|a|^2}, \cdots$, is not less than 1 any more. Therefore, in the geometric slicing, we replace geometric sequences with arithmetic sequences.

• Geometric Slicing: We apply the slicing idea of Lemma 5.2 to get a finite-horizon problem. By putting $\alpha_{k_1} = \frac{1}{k-k_1}$, $\alpha_{k_1+1} = \frac{1}{k-k_1}$, $\cdots$, $\alpha_k = \frac{1}{k-k_1}$ and $\beta_{k_2} = \frac{1}{k-k_2}$, $\beta_{k_2+1} = \frac{1}{k-k_2}$, $\cdots$, $\beta_{k-1} = \frac{1}{k-k_2}$ the average cost is lower bounded by

$$\begin{aligned}&\inf_{u_1,u_2} (q\mathbb{E}[x^2[k]] \\ &+ r_1 \underbrace{(\frac{1}{k-k_1}\mathbb{E}[u_1^2[k_1]] + \cdots + \frac{1}{k-k_1}\mathbb{E}[u_1^2[k-1]])}_{:=\widetilde{P_1}} \\ &+ r_2 \underbrace{(\frac{1}{k-k_2}\mathbb{E}[u_2^2[k_2]] + \cdots + \frac{1}{k-k_2}\mathbb{E}[u_2^2[k-1]])}_{:=\widetilde{P_2}})\end{aligned}$$

• Three stage division: As we did in the proof of Lemma 5.3, we divide the resulting finite-horizon problem into three time intervals — information-limited interval, Witsenhausen's interval, power-limited interval. Define

$$W_1 := w[0] + \cdots + w[k_1 - 2]$$
$$W_2 := w[k_1 - 1] + \cdots + w[k_2 - 2]$$
$$W_3 := w[k_2 - 1] + \cdots + w[k - 2]$$
$$U_{11} := u_1[1] + \cdots + u_1[k_1 - 1]$$
$$U_{21} := u_2[1] + \cdots + u_2[k_1 - 1]$$
$$U_{22} := u_2[k_1] + \cdots + u_2[k_2 - 1]$$
$$U_1 := u_1[k_1] + \cdots + u_1[k - 1]$$
$$U_2 := u_2[k_2] + \cdots + u_2[k - 1]$$
$$X_1 := W_1 + U_{11} + U_{12}$$
$$X_2 := W_2 + U_{22}$$

Like the proof of Lemma 5.3, W_1, W_2, W_3 represent the distortions of three intervals. U_{11} and U_{21} represent the first and second controller inputs in the information-limited interval. U_1 represents the remaining input of the first controller. U_{22} and U_2 represent the second controller's input in Witsenhausen's and power-limited intervals respectively.

The goal of this proof is grouping control inputs, so that we reveal the effects of the controller inputs on the state and isolate their effects according to their characteristics.

• Power-Limited Inputs: We first isolate the power-limited inputs, i.e. the first controller's input in the Witsenhausen's and power-limited interval, and the second controller's input in the power-limited interval. Notice that

$$\begin{aligned}
x[k] &= w[k-1] + w[k-2] + \cdots + w[0] \\
&\quad + u_1[k-1] + u_1[k-2] + \cdots + u_1[0] \\
&\quad + u_2[k-1] + u_2[k-2] + \cdots + u_2[0] \\
&= (w[0] + \cdots + w[k_1 - 2] \\
&\quad + u_1[1] + \cdots + u_1[k_1 - 1] \\
&\quad + u_2[1] + \cdots + u_2[k_1 - 1]) \\
&\quad + (w[k_1 - 1] + \cdots + w[k_2 - 2] \\
&\quad + u_2[k_1] + \cdots + u_2[k_2 - 1]) \\
&\quad + (w[k_2 - 1] + \cdots + w[k-2]) \\
&\quad + (u_1[k_1] + \cdots + u_1[k-1]) \\
&\quad + (u_2[k_2] + \cdots + u_2[k-1]) \\
&\quad + w[k-1].
\end{aligned}$$

Therefore, by Lemma 4.1 of Chapter 4 we have

$$\begin{aligned}
\mathbb{E}[x^2[k]] &= \mathbb{E}[(X_1 + X_2 + W_3 + U_1 + U_2 + w[k-1])^2] \\
&= \mathbb{E}[(X_1 + X_2 + W_3 + U_1 + U_2)^2] + \mathbb{E}[w^2[k-1]] \\
&\geq (\sqrt{\mathbb{E}[(X_1 + X_2 + W_3)^2]} - \sqrt{\mathbb{E}[U_1^2]} - \sqrt{\mathbb{E}[U_2^2]})_+^2 + 1 \\
&= (\sqrt{\mathbb{E}[(X_1 + X_2)^2] + \mathbb{E}[W_3^2]} - \sqrt{\mathbb{E}[U_1^2]} - \sqrt{\mathbb{E}[U_2^2]})_+^2 + 1 \qquad (10.61)
\end{aligned}$$

where the last equality follows from the causality. Here, we can see that $\mathbb{E}[(X_1 + X_2)^2]$ does not depend on the power-limited inputs.

• Information-Limited Interval: We will bound the remaining state distortion after the information-limited interval. Denote y_1' and y_2' as follows:

$$\begin{aligned}
y_1'[k] &= w[0] + w[1] + \cdots + w[k-1] + v_1[k] \\
y_2'[k] &= w[0] + w[1] + \cdots + w[k-1] + v_2[k]
\end{aligned}$$

Here, $y_1'[k]$, $y_2'[k]$ can be obtained by removing $u_1[1:k-1]$, $u_2[1:k-1]$ from $y_1[k]$, $y_2[k]$, and $u_1[k]$ and $u_2[k]$ are functions of $y_1[1:k]$ and $y_2[1:k]$ respectively. Therefore, we can see that $y_1[1:k]$, $y_2[1:k]$ are functions of $y_1'[1:k]$, $y_2'[1:k]$. Moreover, W_1, $y_1'[1:k_1-1]$, $y_2'[1:k_1-1]$ are jointly Gaussian.

Let

$$\begin{aligned}
W_1' &:= W_1 - \mathbb{E}[W_1 | y_1'[1:k_1-1], y_2'[1:k_1-1]] \\
W_1'' &:= \mathbb{E}[W_1 | y_1'[1:k_1-1], y_2'[1:k_1-1]].
\end{aligned}$$

Then, W_1', W_1'', W_2 are independent Gaussian random variables. Moreover, W_1', W_2 are independent from $y_1'[1:k_1-1], y_2'[1:k_1-1]$. W_1'' is a function of $y_1'[1:k_1-1], y_2'[1:k_1-1]$.

Now, let's lower bound $\mathbb{E}[(X_1+X_2)^2]$. Since Gaussian maximizes the entropy, we have

$$\begin{aligned}
&\frac{1}{2}\log(2\pi e\mathbb{E}[(X_1+X_2)^2] \\
&\geq h(X_1+X_2) \\
&\geq h(X_1+X_2|y_1'[1:k_1-1], y_2'[1:k_1-1], y_2[k_1:k_2-1]) \\
&= h(W_1'+W_1''+U_{11}+U_{12}+W_2+U_{22}|y_1'[1:k_1-1], y_2'[1:k_1-1], y_2[k_1:k_2-1]) \\
&= h(W_1'+W_2|y_1'[1:k_1-1], y_2'[1:k_1-1], y_2[k_1:k_2-1])
\end{aligned} \tag{10.62}$$

We will first lower bound the variance of W_1'. Notice that

$$\mathbb{E}[y_1'[k]^2] = \mathbb{E}[w^2[0]] + \cdots + \mathbb{E}[w^2[k-1]] + \mathbb{E}[v_1^2[k]] = k + \sigma_{v1}^2$$

and

$$\mathbb{E}[y_2'[k]^2] = k + \sigma_{v2}^2.$$

Thus, we have

$$\begin{aligned}
&I(W_1; y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
&= h(y_1'[1:k_1-1], y_2'[1:k_1-1]) - h(y_1'[1:k_1-1], y_2'[1:k_1-1]|W_1) \\
&\leq \sum_{1\leq i\leq k_1-1} h(y_1'[i]) + \sum_{1\leq i\leq k_1-1} h(y_2'[i]) - \sum_{1\leq i\leq k_1-1} h(v_1[i]) - \sum_{1\leq i\leq k_1-1} h(v_2[i]) \\
&\leq \sum_{1\leq k\leq k_1-1} \frac{1}{2}\log(\frac{k+\sigma_{v1}^2}{\sigma_{v1}^2}) + \sum_{1\leq k\leq k_1-1} \frac{1}{2}\log(\frac{k+\sigma_{v2}^2}{\sigma_{v2}^2}) \\
&= \frac{1}{2}\log(\prod_{1\leq k\leq k_1-1} \frac{k+\sigma_{v1}^2}{\sigma_{v1}^2}) + \frac{1}{2}\log(\prod_{1\leq k\leq k_1-1} \frac{k+\sigma_{v2}^2}{\sigma_{v2}^2}) \\
&\overset{(A)}{\leq} \frac{k_1-1}{2}\log(\frac{1}{k_1-1}\sum_{1\leq k\leq k_1-1} \frac{k+\sigma_{v1}^2}{\sigma_{v1}^2}) + \frac{k_1-1}{2}\log(\frac{1}{k_1-1}\sum_{1\leq k\leq k_1-1} \frac{k+\sigma_{v2}^2}{\sigma_{v2}^2}) \\
&= \frac{k_1-1}{2}\log(1+\frac{1}{k_1-1}\sum_{1\leq k\leq k_1-1} \frac{k}{\sigma_{v1}^2}) + \frac{k_1-1}{2}\log(1+\frac{1}{k_1-1}\sum_{1\leq k\leq k_1-1} \frac{k}{\sigma_{v2}^2}) \\
&\leq \frac{k_1-1}{2}\log(1+\frac{1}{k_1-1}\sum_{1\leq k\leq k_1-1} \frac{k_1-1}{\sigma_{v1}^2}) + \frac{k_1-1}{2}\log(1+\frac{1}{k_1-1}\sum_{1\leq k\leq k_1-1} \frac{k_1-1}{\sigma_{v2}^2}) \\
&= \frac{k_1-1}{2}\log(1+\frac{1}{(k_1-1)\sigma_{v1}^2}(k_1-1)^2) + \frac{k_1-1}{2}\log(1+\frac{1}{(k_1-1)\sigma_{v2}^2}(k_1-1)^2) \\
&= \frac{k_1-1}{2}\log(1+\frac{k_1-1}{\sigma_{v1}^2}) + \frac{k_1-1}{2}\log(1+\frac{k_1-1}{\sigma_{v2}^2})
\end{aligned} \tag{10.63}$$

(A): Arithmetic-Geometric mean.

Let's denote the last equation as I. We also have

$$\mathbb{E}[W_1^2] = k_1 - 1 \tag{10.64}$$

Now, we can bound the variance of the Gaussian random variable W_1' as follows:

$$\begin{aligned}
&\frac{1}{2}\log(2\pi e\mathbb{E}[W_1'^2]) = h(W_1') \\
&\geq h(W_1'|y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
&= h(W_1|y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
&= h(W_1) - I(W_1; y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
&\geq \frac{1}{2}\log(2\pi e(k_1-1)) - I
\end{aligned}$$

where the last inequality follows from (10.63) and (10.64).

Thus,

$$\mathbb{E}[W_1'^2] \geq \frac{k_1-1}{2^{2I}} \tag{10.65}$$

and denote the last term as Σ. Since W_1' is Gaussian, $W_1' = W_1''' + W_1''''$ where $W_1''' \sim \mathcal{N}(0, \Sigma)$, and W_1''', W_1'''' are independent.

Moreover, we also have

$$\mathbb{E}[W_2^2] = \mathbb{E}[(w[k_1-1] + \cdots + w[k_2-2])^2] = k_2 - k_1. \tag{10.66}$$

By (10.62), we have

$$\begin{aligned}
&\frac{1}{2}\log(2\pi e\mathbb{E}[(X_1+X_2)^2]) \\
&\geq h(W_1' + W_2|y_1'[1:k_1-1], y_2'[1:k_1-1], y_2[k_1:k_2-1]) \\
&\geq h(W_1' + W_2|W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1], y_2[k_1:k_2-1]) \\
&= h(W_1''' + W_2|W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1], y_2[k_1:k_2-1]) \\
&= h(W_1''' + W_2|W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
&- I(W_1''' + W_2; y_2[k_1:k_2-1]|W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
&= h(W_1''' + W_2) \\
&- I(W_1''' + W_2; y_2[k_1:k_2-1]|W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
&\geq \frac{1}{2}\log(2\pi e(\Sigma + k_2 - k_1)) \\
&- I(W_1''' + W_2; y_2[k_1:k_2-1]|W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1])
\end{aligned} \tag{10.67}$$

where the last inequality comes from the fact that W_1''' and W_2 are independent Gaussian, and (10.65), (10.66). Now, the question boils down to the upper bound of the last mutual information term, which can be understood as the information contained in the second controller's observation in Witsenhausen's interval.

• Second controller's observation in Witsenhausen's interval: We will bound the amount of information contained in the second controller's observation in Witsenhausen's interval. For $n \geq k_1$, define

$$\begin{aligned} y_2''[n] &:= W_1''' + w[k_1-1] + w[k_1] + \cdots + w[n-1] \\ &+ u_1[k_1] + \cdots + u_1[n-1] \\ &+ v_2[n] \end{aligned}$$

Notice the relationship between $y_2[n]$ and $y_2''[n]$:

$$y_2[n] = y_2''[n] + u_2[k_1] + \cdots + u_2[n-1] + W_1''' + \mathbb{E}[W_1 | y_1'[1:k_1-1], y_2'[1:k_1-1]]. \quad (10.68)$$

The mutual information of (10.67) is bounded as follows:

$$\begin{aligned} & I(W_1''' + W_2; y_2[k_1:k_2-1] | W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\ & = h(y_2[k_1:k_2-1] | W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\ & - h(y_2[k_1:k_2-1] | W_1''' + W_2, W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\ & = \sum_{k_1 \leq i \leq k_2-1} h(y_2[i] | y_2[k_1:i-1], W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\ & - \sum_{k_1 \leq i \leq k_2-1} h(y_2[i] | y_2[k_1:i-1], W_1''' + W_2, W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\ & \overset{(A)}{=} \sum_{k_1 \leq i \leq k_2-1} h(y_2''[i] | y_2[k_1:i-1], W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\ & - \sum_{k_1 \leq i \leq k_2-1} h(y_2[i] | y_2[k_1:i-1], W_1''' + W_2, W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\ & \overset{(B)}{\leq} \sum_{k_1 \leq i \leq k_2-1} h(y_2''[i]) - \sum_{k_1 \leq i \leq k_2-1} h(v_2[i]) \\ & \leq \sum_{k_1 \leq i \leq k_2-1} \frac{1}{2} \log(2\pi e \mathbb{E}[y_2''[i]^2]) - \sum_{k_1 \leq i \leq k_2-1} \frac{1}{2} \log(2\pi e \sigma_{v2}^2) \end{aligned} \quad (10.69)$$

(A): Since $y_2[1:k_1-1]$ is a function of $y_2'[1:k_1-1]$, $u_2[k_1], \cdots, u_2[i]$ are functions of $y_2[k_1:i-1], y_2'[1:k_1-1]$. Thus, all the terms in (10.68) except $y_2''[i]$ can be vanished by the conditioning
(B): By causality, $v_2[i]$ is independent from all conditioning random variables.

First, let's bound the variance of $y_2''[n]$. By Lemma 4.1 of Chapter 4, we have

$$\begin{aligned} \mathbb{E}[y_2''[n]^2] &\leq 2\mathbb{E}[(W_1''' + w[k_1-1] + w[k_1] + \cdots + w[n-1])^2] \\ &+ 2\mathbb{E}[(u_1[k_1] + \cdots + u_1[n-1])^2] + \sigma_{v2}^2 \\ &= 2(\Sigma + n - k_1 + 1) + 2\mathbb{E}[(u_1[k_1] + \cdots + u_1[n-1])^2] + \sigma_{v2}^2. \end{aligned}$$

By Lemma 4.1 of Chapter 4, we have

$$\begin{aligned}
&\mathbb{E}[(u_1[k_1] + \cdots + u_1[n-1])^2] \\
&\leq (\sqrt{\mathbb{E}[u_1^2[k_1]]} + \cdots + \sqrt{\mathbb{E}[u_1^2[n-1]]})^2 \\
&\overset{(A)}{\leq} (n-k_1)(\mathbb{E}[u_1^2[k_1]] + \mathbb{E}[u_1^2[k_1+1]] + \cdots + \mathbb{E}[u_1^2[n-1]]) \\
&\leq (n-k_1)(k-k_1)\widetilde{P_1}.
\end{aligned}$$

(A): Cauchy-Schwarz inequality

Thus, the variance of $y_2''[n]$ is bounded as:

$$\mathbb{E}[y_2''[n]^2] \leq 2\Sigma + 2(n-k_1+1) + 2(n-k_1)(k-k_1)\widetilde{P_1} + \sigma_{v2}^2$$

Therefore, we have

$$\begin{aligned}
&\sum_{k_1 \leq n \leq k_2-1} \mathbb{E}[y_2''[n]^2] \\
&\leq \sum_{k_1 \leq n \leq k_2-1} (2\Sigma + 2(n-k_1+1) + 2(n-k_1)(k-k_1)\widetilde{P_1} + \sigma_{v2}^2) \\
&\leq 2(k_2-k_1)\Sigma + \sum_{k_1 \leq n \leq k_2-1} 2(k_2-k_1) + \sum_{k_1 \leq n \leq k_2-1} 2(k_2-k_1-1)(k-k_1)\widetilde{P_1} + (k_2-k_1)\sigma_{v2}^2 \\
&= 2(k_2-k_1)\Sigma + 2(k_2-k_1)^2 + 2(k_2-k_1)(k_2-k_1-1)(k-k_1)\widetilde{P_1} + (k_2-k_1)\sigma_{v2}^2 \qquad (10.70)
\end{aligned}$$

Therefore, by (10.69) and (10.70) we conclude

$$\begin{aligned}
&I(W_1''' + W_2; y_2[k_1 : k_2-1] | W'''', y_1'[1 : k_1-1], y_2'[1 : k_1-1]) \\
&\leq \sum_{k_1 \leq n \leq k_2-1} \frac{1}{2} \log(\frac{\mathbb{E}[y_2''[n]^2]}{\sigma_{v2}^2}) \\
&= \frac{1}{2} \log(\prod_{k_1 \leq n \leq k_2-1} \frac{\mathbb{E}[y_2''[n]^2]}{\sigma_{v2}^2}) \\
&\overset{(A)}{\leq} \frac{k_2-k_1}{2} \log(\frac{1}{k_2-k_1} \sum_{k_1 \leq n \leq k_2-1} \frac{\mathbb{E}[y_2''[n]^2]}{\sigma_{v2}^2}) \\
&\leq \frac{k_2-k_1}{2} \log(1 + \frac{1}{(k_2-k_1)\sigma_{v2}^2}(2(k_2-k_1)\Sigma + 2(k_2-k_1)^2 + 2(k_2-k_1)(k_2-k_1-1)(k-k_1)\widetilde{P_1})) \\
&\leq \frac{k_2-k_1}{2} \log(1 + \frac{1}{\sigma_{v2}^2}(2\Sigma + 2(k_2-k_1) + 2(k_2-k_1-1)(k-k_1)\widetilde{P_1}))
\end{aligned}$$

(A): Arithmetic-Geometric mean

Denote the last equation as $I'(\widetilde{P_1})$. By (10.67), we can conclude

$$\frac{1}{2}\log(2\pi e \mathbb{E}[(X_1+X_2)^2] \geq \frac{1}{2}\log(2\pi e(\Sigma + k_2 - k_1)) - I'(\widetilde{P_1})$$

which implies

$$\mathbb{E}[(X_1+X_2)^2] \geq \frac{\Sigma + k_2 - k_1}{2^{2I'(\widetilde{P_1})}}. \qquad (10.71)$$

• Final lower bound: Now, we can merge the inequalities to prove the lemma. The variance of W_3 is

$$\mathbb{E}[W_3^2] = k - k_2. \tag{10.72}$$

By Lemma 4.1 of Chapter 4 and Cauchy-Schwarz inequality, the variance of U_1 is upper bounded as follows:

$$\begin{aligned}\mathbb{E}[U_1^2] &\leq (\sqrt{a^{2(k-k_1-1)}\mathbb{E}[u_1^2[k_1]]} + \cdots + \sqrt{\mathbb{E}[u_1^2[k-1]]})^2 \\ &\leq (k-k_1)(\mathbb{E}[u_1^2[k_1]] + \mathbb{E}[u_1^2[k_1+1]] + \cdots + \mathbb{E}[u_1^2[k-1]]) \\ &= (k-k_1)^2\widetilde{P_1}. \end{aligned} \tag{10.73}$$

Likewise, the variance of U_2 can be bounded as

$$\mathbb{E}[U_2^2] \leq (k-k_2)^2\widetilde{P_2}. \tag{10.74}$$

By plugging (10.71), (10.72), (10.73), (10.74) into (10.61), we finally prove the lemma. □

Proof of Corollary 5.5 of Page 250. Proof of (a):

Since $\sigma_{v2} \geq 16$, we can find $k_2 \geq 6$ such that

$$k_2 - 2 \leq \frac{\sigma_{v2}}{4} < k_2 - 1 \tag{10.75}$$

We put such k_3, $k_1 = 1$ and $k = k_2$ as the parameters of Lemma 5.4. Then, the lower bound of Lemma 5.4 reduces to

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq (\sqrt{\frac{k_2-1}{2^{2I'(\widetilde{P_1})}}} - \sqrt{(k_2-1)^2\widetilde{P_1}})_+^2 + 1. \tag{10.76}$$

Since $k_2 \geq 6$, we have

$$\frac{k_2-2}{k_2-1} \geq \frac{4}{5}. \tag{10.77}$$

Thus, $I'(\widetilde{P_1})$ is lower bounded by

$$\begin{aligned}I'(\widetilde{P_1}) &= \frac{1}{2}\log(1 + \frac{1}{\sigma_{v2}^2}(2(k_2-1) + 2(k_2-2)(k_2-1)\widetilde{P_1}))^{k_2-1} \\ &= \frac{1}{2}\log(1 + \frac{1}{k_2-1}(\frac{2(k_2-1)^2}{\sigma_{v2}^2} + \frac{2(k_2-2)(k_2-1)^2\widetilde{P_1}}{\sigma_{v2}^2}))^{k_2-1} \\ &\overset{(A)}{\leq} \frac{1}{2}\log(1 + \frac{1}{k_2-1}(2(\frac{5}{4})^2\frac{(k_2-2)^2}{\sigma_{v2}^2} + 2(\frac{5}{4})^2\frac{(k_2-2)(k_2-2)^2}{4\sigma_{v2}^3}))^{k_2-1} \\ &\overset{(B)}{\leq} \frac{1}{2}\log(1 + \frac{1}{k_2-1}(2(\frac{5}{4})^2(\frac{1}{4})^2 + 2(\frac{5}{4})^2(\frac{1}{4})^3))^{k_2-1} \\ &\leq \frac{1}{2}\log e^{\frac{125}{512}} \end{aligned} \tag{10.78}$$

(A): (10.77) and $\widetilde{P_1} \leq \frac{1}{4\sigma_{v2}}$.
(B): (10.75).

Moreover, we have

$$\begin{aligned}(k_2-1)^2\widetilde{P_1} &\overset{(A)}{\leq} \frac{5}{4}(k_2-1)(k_2-2)\widetilde{P_1} \\ &\overset{(B)}{\leq} \frac{5}{4}(k_2-1)\frac{k_2-2}{4\sigma_{v2}} \\ &\overset{(C)}{\leq} \frac{5}{4}(k_2-1)\frac{1}{16} \\ &= \frac{5}{64}(k_2-1) \end{aligned} \tag{10.79}$$

(A): (10.77)
(B): $\widetilde{P_1} \leq \frac{1}{4\sigma_{v2}}$
(C): (10.75)

Therefore, by plugging (10.78), (10.79) into (10.76), we get

$$\begin{aligned} D_L(\widetilde{P_1},\widetilde{P_2}) &\geq (\sqrt{\frac{k_2-1}{2^{2I'(P_1)}}} - \sqrt{(k_2-1)^2\widetilde{P_1}})_+^2 + 1 \\ &\geq (\sqrt{\frac{k_2-1}{e^{\frac{125}{512}}}} - \sqrt{\frac{5}{64}(k_2-1)})_+^2 + 1 \\ &= 0.366724...(k_2-1)+1 \\ &\geq 0.366724...\frac{\sigma_{v2}}{4}+1 \\ &= 0.09168106...\sigma_{v2}+1 \end{aligned}$$

where the last inequality follows from (10.75).

Proof of (b):

Since $\frac{1}{\widetilde{P_1}} \geq 64$, we can find $k_2 \geq 6$ such that

$$k_2 - 2 \leq \frac{1}{16\widetilde{P_1}} < k_2 - 1. \tag{10.80}$$

We put such k_2, $k_1 = 1$ and $k = k_2$ as the parameters of Lemma 5.4. Then, the lower bound of Lemma 5.4 reduces to

$$D_L(\widetilde{P_1},\widetilde{P_2}) \geq (\sqrt{\frac{k_2-1}{2^{2I'(\widetilde{P_1})}}} - \sqrt{(k_2-1)^2\widetilde{P_1}})_+^2 + 1. \tag{10.81}$$

First, $I'(\widetilde{P_1})$ is lower bounded by

$$\begin{aligned}
I'(\widetilde{P_1}) &= \frac{1}{2}\log(1+\frac{1}{\sigma_{v2}^2}(2(k_2-1)+2(k_2-2)(k_2-1)\widetilde{P_1}))^{k_2-1} \\
&= \frac{1}{2}\log(1+\frac{1}{k_2-1}(\frac{2(k_2-1)^2}{\sigma_{v2}^2}+\frac{2(k_2-2)(k_2-1)^2\widetilde{P_1}}{\sigma_{v2}^2}))^{k_2-1} \\
&\overset{(A)}{\le} \frac{1}{2}\log(1+\frac{1}{k_2-1}(2(k_2-1)^2(4\widetilde{P_1})^2+2(k_2-2)(k_2-1)^2\widetilde{P_1}(4\widetilde{P_1})^2))^{k_2-1} \\
&\overset{(B)}{\le} \frac{1}{2}\log(1+\frac{1}{k_2-1}(2(\frac{5}{4})^2(k_2-2)^2(4\widetilde{P_1})^2+2(\frac{5}{4})^2(k_2-2)^3\widetilde{P_1}(4\widetilde{P_1})^2))^{k_2-1} \\
&\overset{(C)}{\le} \frac{1}{2}\log(1+\frac{1}{k_2-1}(2(\frac{5}{4})^2(\frac{1}{4})^2+2(\frac{5}{4})^2\frac{1}{16}(\frac{1}{4})^2)^{k_2-1} \\
&\le \frac{1}{2}\log e^{\frac{425}{2048}}.
\end{aligned} \tag{10.82}$$

(A): $\frac{1}{4\sigma_{v2}} \le P_1$

(B): Since $k_2 \ge 6$, (10.77) still holds.

(C): (10.80)

Moreover, we also have

$$\begin{aligned}
(k_2-1)^2\widetilde{P_1} &\overset{(A)}{\le} \frac{5}{4}(k_2-1)(k_2-2)\widetilde{P_1} \\
&\overset{(B)}{\le} \frac{5}{4}(k_2-1)\frac{1}{16} \\
&= \frac{5}{64}(k_2-1).
\end{aligned} \tag{10.83}$$

(A): Since $k_2 \ge 6$, (10.77) still holds.

(B): (10.80)

Therefore, plugging (10.82), (10.83) into (10.81) we can conclude

$$\begin{aligned}
D_L(\widetilde{P_1},\widetilde{P_2}) &\ge (\sqrt{\frac{k_2-1}{2^{2I'(\widetilde{P_1})}}}-\sqrt{(k_2-1)^2\widetilde{P_1}})_+^2+1 \\
&\ge (\sqrt{\frac{k_2-1}{e^{\frac{425}{2048}}}}-\sqrt{\frac{5}{64}(k_2-1)})_+^2+1 \\
&= 0.386801...(k_2-1)+1 \\
&\ge 0.386801...\frac{1}{16\widetilde{P_1}}+1 \\
&= \frac{0.0241750...}{16\widetilde{P_1}}+1
\end{aligned}$$

where the last inequality comes from (10.80).

Proof of (c):

Denote $P := \max(\widetilde{P_1},\widetilde{P_2})$. Since $P \le \frac{1}{50}$, there exists $k \ge 3$ such that

$$k-2 \le \frac{1}{50P} < k-1. \tag{10.84}$$

We put such k and $k_1 = k_2 = 1$ as the parameters of Lemma 5.4. Then, the lower bound of Lemma 5.4 reduces to

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq (\sqrt{k-1} - \sqrt{(k-1)^2\widetilde{P_1}} - \sqrt{(k-1)^2\widetilde{P_2}})_+^2 + 1.$$

Since $k \geq 3$, we have

$$\frac{k-2}{k-1} \geq \frac{1}{2}. \tag{10.85}$$

Therefore, we conclude

$$\begin{aligned} D_L(\widetilde{P_1}, \widetilde{P_2}) &\geq (\sqrt{k-1} - \sqrt{4(k-1)^2P})_+^2 + 1 \\ &\overset{(A)}{\geq} (\sqrt{k-1} - \sqrt{16(k-2)^2P})_+^2 + 1 \\ &\overset{(B)}{\geq} (\sqrt{\frac{1}{50P}} - \sqrt{\frac{16}{50^2P}})_+^2 + 1 \\ &\geq 0.00377258...\frac{1}{P} + 1. \end{aligned}$$

(A): (10.84)
(B): (10.85)

Proof of (d):

As mentioned in the proof of Corollary 5.4 (j), the centralized controller's distortion that has both observations $y_1[n], y_2[n]$ and has no input power constraints is a lower bound on the decentralized controller's distortion.

Let $y_1'[n] := x[n] + v_1'[n]$ and $y_2'[n] := x[n] + v_2'[n]$ where $v_1'[n] \sim \mathcal{N}(0, \sigma_1^2)$ and $v_2'[n] \sim \mathcal{N}(0, \sigma_1^2)$ are i.i.d. random variables. Just like the proof of Corollary 5.4 (j), the performance of the centralized controller with both observations is equivalent to a centralized controller with observation $\frac{y_1'[n]+y_2'[n]}{2}$ by the maximum ratio combining.

Let Σ_E be the estimation error of the Kalman filtering with a scalar observation $\frac{y_1'[n]+y_2'[n]}{2}$. By Lemma 5.1,

$$\begin{aligned} \Sigma_E &= \frac{-1 + \sqrt{4\frac{\sigma_{v1}^2}{2} + 1}}{2} \\ &= \frac{-1 + \sqrt{2\sigma_{v1}^2 + 1}}{2}. \end{aligned}$$

Then, for all $\widetilde{P_1}$ and $\widetilde{P_2}$, the cost of the decentralized controllers is lower bounded as follows:

$$\begin{aligned} D_L(\widetilde{P_1}, \widetilde{P_2}) &\overset{(A)}{\geq} \inf_{|1-k|<1} \frac{(2k-k^2)\Sigma_E + 1}{1-(1-k)^2} \\ &= \inf_{|1-k|<1} \Sigma_E + \frac{1}{1-(1-k)^2} \\ &\geq \Sigma_E. \end{aligned}$$

(A): The decentralized control cost is larger than the centralized controller's cost with the observation $\frac{y_1'[n]+y_2'[n]}{2}$. Moreover, when $|a-k| \geq 1$ the centralized control system is unstable, and the cost diverges to infinity. When $|a-k| < 1$, the cost analysis follows from Lemma 5.1.

By Lemma 5.4, $D_L(\widetilde{P_1}, \widetilde{P_2}) \geq 1$. Finally, for all $\widetilde{P_1}, \widetilde{P_2}$ we have

$$\begin{aligned} D_L(\widetilde{P_1}, \widetilde{P_2}) &\geq \max(\Sigma_E, 1) = \max(\frac{-1+\sqrt{2\sigma_{v1}^2+1}}{2}, 1) \\ &\geq \frac{1}{2}(\frac{-1+\sqrt{2\sigma_{v1}^2+1}}{2}) + \frac{1}{2} \\ &\geq \frac{1}{4} + \frac{\sqrt{2\sigma_{v1}^2+1}}{2} \\ &\geq \frac{\sqrt{2}}{2}\sigma_{v1}. \end{aligned}$$

Since we already know $D_L(\widetilde{P_1}, \widetilde{P_2}) \geq 1$, the statement (d) of the corollary is true. □

Proof of Proposition 5.2 of Page 250. Like the proof of Proposition 5.1, we define the subscript max as $argmax_{i \in \{1,2\}} \widetilde{P_i}$, and write $D_{\sigma_{v1}}(\cdot), D_{\sigma_{v2}}(\cdot), D_{\sigma_{vmax}}(\cdot)$ as $D_{v1}(\cdot), D_{v2}(\cdot), D_{vmax}(\cdot)$ respectively.

By the same argument as the proof of Proposition 5.1, it is enough to show that there exists $c \leq 10^6$ such that for all $\widetilde{P_1}, \widetilde{P_2} \geq 0$, $\min(D_{\sigma 1}(c\widetilde{P_1}), D_{\sigma 2}(c\widetilde{P_2})) \leq c \cdot D_L(\widetilde{P_1}, \widetilde{P_2})$.

In the proof, we first divide the cases based on σ_1, σ_2 (essentially equivalent to Σ_1, Σ_2), and then based on $\widetilde{P_1}, \widetilde{P_2}$. Here, we can use the fact that $\sigma_1 \leq \sigma_2$ to reduce the cases.

(i) When $\sigma_{v1} \leq 16$, $\sigma_{v2} \leq 16$

(i-i) If $\max(\widetilde{P_1}, \widetilde{P_2}) \geq \frac{1}{64}$

Lower bound: By Corollary 5.5 (d),

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq 1.$$

Upper bound: Since $\sigma_{v1}, \sigma_{v2} \leq 16$, we can plug $t = \frac{1}{15.008}$ into the equation (5.9) of Corollary 5.1. Thus, we have

$$(D_{\sigma max}(P_{max}), P_{max}) \leq (30.016, \frac{1}{15.008}).$$

Ratio: c is upper bounded by

$$c \leq 30.016.$$

(i-ii) If $\max(\widetilde{P_1}, \widetilde{P_2}) \leq \frac{1}{64}$

Lower bound: Since $\max(\widetilde{P_1}, \widetilde{P_2}) \leq \frac{1}{64} \leq \frac{1}{50}$, by Corollary 5.5 (c) we can conclude

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq \frac{0.003772}{\max(\widetilde{P_1}, \widetilde{P_2})} + 1.$$

Upper bound: Since $\sigma_{v1}, \sigma_{v2} \leq 16$ and $\widetilde{P_{max}} \leq \frac{1}{64} \leq \frac{1}{15.008}$, we can plug $t = \widetilde{P_{max}}$ into the equation (5.9) of Corollary 5.1. Thus, we have

$$(D_{\sigma max}(P_{max}), P_{max}) \leq (\frac{2}{\widetilde{P_{max}}}, \widetilde{P_{max}}).$$

Ratio: c is upper bounded by

$$c \leq \frac{2}{0.003772} < 540.$$

(ii) When $\sigma_{v1} \leq 16 \leq \sigma_{v2}$

(ii-i) If $\widetilde{P_1} \geq \frac{1}{64}$

Lower bound: By Corollary 5.5 (d), we have

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq 1.$$

Upper bound: Since $\sigma_1 \leq 16$, we can plug $t = \frac{1}{15.008}$ into the equation (5.9) of Corollary 5.1. Thus, we have

$$(D_{\sigma 1}(P_1), P_1) \leq (30.016, \frac{1}{15.008}).$$

Ratio: c is upper bounded by

$$c \leq 30.016.$$

(ii-ii) If $\frac{1}{4\sigma_{v2}} \leq \widetilde{P_1} \leq \frac{1}{64}$

Lower bound: By Corollary 5.5 (b), we have

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq \frac{0.02417}{\widetilde{P_1}} + 1.$$

Upper bound: Since $\widetilde{P_1} \leq \frac{1}{64} \leq \frac{1}{15.008}$, we can plug $t = \widetilde{P_1}$ into the equation (5.9) of Corollary 5.1. Thus, we have

$$(D_{\sigma 1}(P_1), P_1) \leq (\frac{2}{\widetilde{P_1}}, \widetilde{P_1}).$$

Ratio: c is upper bounded by

$$c \leq \frac{2}{0.02417} < 83.$$

(ii-iii) If $\widetilde{P_1} \leq \frac{1}{4\sigma_{v2}}$ and $\widetilde{P_2} \geq \frac{1}{4\sigma_{v2}}$

Lower bound: By Corollary 5.5 (a), we have

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq 0.09168\sigma_{v2} + 1$$

Upper bound: Since $\sigma_{v2} \geq 16$, we can plug $t = \frac{1}{1.0005\sigma_{v2}}$ into the equation (5.8) of Corollary 5.1. Thus, we have

$$(D_{\sigma 2}(\widetilde{P_2}), \widetilde{P_2}) \leq (2.001\sigma_{v2}, \frac{1}{1.0005\sigma_{v2}}).$$

Ratio: c is upper bounded by

$$c \leq \frac{2.001}{0.09168} < 22.$$

(ii-iv) If $\widetilde{P_1} \le \frac{1}{4\sigma_{v2}}$ and $\widetilde{P_2} \le \frac{1}{4\sigma_{v2}}$

Lower bound: Since $\widetilde{P_1} \le \frac{1}{4\sigma_{v2}} \le \frac{1}{64} \le \frac{1}{50}$, $\widetilde{P_2} \le \frac{1}{4\sigma_{v2}} \le \frac{1}{64} \le \frac{1}{50}$, by Corollary 5.5 (c) we have

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \ge \frac{0.003772}{\max(\widetilde{P_1}, \widetilde{P_2})} + 1.$$

Upper bound: Since $\widetilde{P_1} \le \frac{1}{4\sigma_{v2}} \le \frac{1}{64} \le \frac{1}{15.008}$, $\widetilde{P_2} \le \frac{1}{4\sigma_{v2}} \le \frac{1}{1.0005\sigma_{v2}}$, these satisfies the conditions for (5.8), (5.9) of Corollary 5.1 respectively. Therefore, by plugging $t = \widetilde{P_{max}}$ into Corollary 5.1, we have

$$(D_{\sigma max}(P_{max}), P_{max}) \le (\frac{2}{\widetilde{P_{max}}}, \widetilde{P_{max}}).$$

Ratio: c is upper bounded by

$$c \le \frac{2}{0.003772} < 540.$$

(iii) When $\sigma_{v1} \ge 16$ and $\sigma_{v2} \ge 16$

(iii-i) If $\widetilde{P_1} \ge \frac{1}{4\sigma_{v1}}$

Lower bound: By Corollary 5.5 (d), we have

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \ge \frac{\sqrt{2}}{2}\sigma_{v1}.$$

Upper bound: Since $\sigma_{v1} \ge 16$, we can plug $t = \frac{1}{1.0005\sigma_{v1}}$ into (5.8) of Corollary 5.1. Thus, we have

$$(D_{\sigma 1}(P_1), P_1) \le (2.001\sigma_{v1}, \frac{1}{1.0005\sigma_{v1}}).$$

Ratio: c is upper bounded by

$$c \le \frac{4}{1.0005} < 4.$$

(iii-ii) If $\frac{1}{4\sigma_{v2}} \le \widetilde{P_1} \le \frac{1}{4\sigma_{v1}}$

Lower bound: Since $\frac{1}{4\sigma_{v2}} \le \widetilde{P_1} \le \frac{1}{4\sigma_{v1}} \le \frac{1}{64}$, by Corollary 5.5 (b) we have

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \ge \frac{0.02417}{\widetilde{P_1}} + 1.$$

Upper bound: Since $\frac{1}{\widetilde{P_1}} \le \frac{1}{4\sigma_{v1}} \le \frac{1}{1.0005\sigma_{v1}}$, we can plug $t = \widetilde{P_1}$ into the equation (5.8) of Corollary 5.1. Thus, we have

$$(D_{\sigma 1}(P_1), P_1) \le (\frac{2}{\widetilde{P_1}}, \widetilde{P_1}).$$

Ratio: c is upper bounded by

$$c \le \frac{2}{0.02417} < 83.$$

(iii-iii) If $\widetilde{P_1} \le \frac{1}{4\sigma_{v2}}$ and $\widetilde{P_2} \ge \frac{1}{4\sigma_{v2}}$

Lower bound: By Corollary 5.5 (a), we have

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \ge 0.09168\sigma_{v2} + 1.$$

Upper bound: Since $\sigma_{v2} \ge 16$, we can plug $t = \frac{1}{1.0005\sigma_{v2}}$ into the equation (5.8) of Corollary 5.1. Thus, we have

$$(D_{\sigma 2}(P_2), P_2) \le (2.001\sigma_{v2}, \frac{1}{1.0005\sigma_{v2}}).$$

Ratio: c is upper bounded by

$$c \le \frac{2.001}{0.09168} < 22.$$

(iii-iv) If $\widetilde{P_1} \le \frac{1}{4\sigma_{v2}}$ and $\widetilde{P_2} \le \frac{1}{4\sigma_{v2}}$

Lower bound: Since $\widetilde{P_1} \le \frac{1}{4\sigma_{v2}} \le \frac{1}{64} \le \frac{1}{50}$, $\widetilde{P_2} \le \frac{1}{4\sigma_{v2}} \le \frac{1}{64} \le \frac{1}{50}$, by Corollary 5.5 (c) we have

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \ge \frac{0.003772}{\max(\widetilde{P_1}, \widetilde{P_2})} + 1.$$

Upper bound: Since $\widetilde{P_1} \le \frac{1}{4\sigma_{v2}} \le \frac{1}{4\sigma_{v1}} \le \frac{1}{1.0005\sigma_{v1}}$ and $\widetilde{P_2} \le \frac{1}{4\sigma_{v2}} \le \frac{1}{1.0005\sigma_{v2}}$, we can plug $t = \widetilde{P_{max}}$ into the equation (5.8) of Corollary 5.1. Thus, we have

$$(D_{\sigma max}(P_{max}), P_{max}) \le (\frac{2}{\widetilde{P_{max}}}, \widetilde{P_{max}}).$$

Ratio: c is upper bounded by

$$c \le \frac{2}{0.003772} < 540.$$

Finally, by (i), (ii), (iii), the constant c is upper bounded by 10^6 and the proposition is proved. □

10.4 Proof of Lemma 5.5, Corollary 5.1 and Proposition 5.2

Proof of Lemma 5.5 of Page 251. For simplicity, we assume $0 \le a < 1$, $1 < k_1 < k_2 < k$. The remaining case when $-1 < a \le 0$ or $k_1 = 1$ or $k_1 = k_2$ or $k = k_2$ easily follow with minor modifications.

• Geometric Slicing: We apply the geometric slicing idea of Lemma 5.2 to get a finite-horizon problem. By putting $\alpha_{k_1} = (\frac{1-a}{1-a^{k-k_1}})a^{k-k_1-1}$, $\alpha_{k_1+1} = (\frac{1-a}{1-a^{k-k_1}})a^{k-k_1-2}$, $\cdots$, $\alpha_k = \frac{1-a}{1-a^{k-k_1}}$ and $\beta_{k_2} = (\frac{1-a^{-1}}{1-a^{-(k-k_2)}})$, $\beta_{k_2+1} = (\frac{1-a^{-1}}{1-a^{-(k-k_2)}})a^{-1}$, $\cdots$, $\beta_{k-1} = (\frac{1-a^{-1}}{1-a^{-(k-k_2)}})a^{-k+1+k_2}$ the average cost is lower bounded by

$$q\mathbb{E}[x^2[k]]$$
$$+ r_1 \underbrace{((\frac{1-a}{1-a^{k-k_1}})a^{k-k_1-1}\mathbb{E}[u_1^2[k_1]] + (\frac{1-a}{1-a^{k-k_1}})a^{k-k_1-2}\mathbb{E}[u_1^2[k_1+1]] + \cdots + (\frac{1-a}{1-a^{k-k_1}})\mathbb{E}[u_1^2[k-1]])}_{:=\widetilde{P_1}}$$

$$+ r_2 \underbrace{((\frac{1-a}{1-a^{k-k_2}})a^{k-k_1-1}\mathbb{E}[u_2^2[k_2]] + (\frac{1-a}{1-a^{k-k_2}})a^{k-k_1-2}\mathbb{E}[u_2^2[k_2+1]] + \cdots + (\frac{1-a}{1-a^{k-k_2}})\mathbb{E}[u_2^2[k-1]])}_{:=\widetilde{P_2}})$$

Here, we denote the second and third terms as $\widetilde{P_1}$ and $\widetilde{P_2}$ respectively.

• Three stage division: As we did in the proof of Lemma 5.3, we will divide the finite-horizon problem into three time intervals — information-limited interval, Witsenhausen's interval, power-limited interval. Define

$$\begin{aligned}
W_1 &:= a^{k-1}w[0] + \cdots + a^{k-k_1+1}w[k_1-2] \\
W_2 &:= a^{k-k_1}w[k_1-1] + \cdots + a^{k-k_2+1}w[k_2-2] \\
W_3 &:= a^{k-k_2}w[k_2-1] + \cdots + aw[k-2] \\
U_{11} &:= a^{k-2}u_1[1] + \cdots + a^{k-k_1}u_1[k_1-1] \\
U_{21} &:= a^{k-2}u_2[1] + \cdots + a^{k-k_1}u_2[k_1-1] \\
U_1 &:= a^{k-k_1-1}u_1[k_1] + \cdots + u_1[k-1] \\
U_{22} &:= a^{k-k_1-1}u_2[k_1] + \cdots + a^{k-k_2}u_2[k_2-1]) \\
U_2 &:= a^{k-k_2-1}u_2[k_2] + \cdots + u_2[k-1] \\
X_1 &:= W_1 + U_{11} + U_{21} \\
X_2 &:= W_2 + U_{22}
\end{aligned}$$

W_1, W_2, W_3 represent the distortions of three intervals respectively. U_{11} and U_{21} represent the first and second controller inputs in the information-limited interval respectively. U_1 represent the remaining input of the first controller. U_{22} and U_2 represent the second controller's input in Witsenhausen's and power-limited intervals respectively.

The goal of this proof is grouping control inputs and expanding $x[n]$, so that we reveal the effects of the controller inputs on the state and isolate their effects according to their characteristics.

• Power-Limited Inputs: We will first isolate the power limited inputs, i.e. the first controller's input in the Witsenhausen's and power-limited intervals, and the second controller's input in the power-limited interval. Notice that

$$\begin{aligned}
x[k] &= w[k-1] + aw[k-2] + \cdots + a^{k-1}w[0] \\
&\quad + u_1[k-1] + au_1[k-2] + \cdots + a^{k-1}u_1[0] \\
&\quad + u_2[k-1] + au_2[k-2] + \cdots + a^{k-1}u_2[0] \\
&= (a^{k-1}w[0] + \cdots + a^{k-k_1+1}w[k_1-2] \\
&\quad + a^{k-2}u_1[1] + \cdots + a^{k-k_1}u_1[k_1-1] \\
&\quad + a^{k-2}u_2[1] + \cdots + a^{k-k_1}u_2[k_1-1]) \\
&\quad + (a^{k-k_1}w[k_1-1] + \cdots + a^{k-k_2+1}w[k_2-2] \\
&\quad + a^{k-k_1-1}u_2[k_1] + \cdots + a^{k-k_2}u_2[k_2-1]) \\
&\quad + (a^{k-k_2}w[k_2-1] + \cdots + aw[k-2]) \\
&\quad + (a^{k-k_1-1}u_1[k_1] + \cdots + u_1[k-1]) \\
&\quad + (a^{k-k_2-1}u_2[k_2] + \cdots + u_2[k-1]) \\
&\quad + w[k-1].
\end{aligned}$$

Therefore, by Lemma 4.1 of Chapter 4 we have

$$\begin{aligned}
\mathbb{E}[x^2[k]] &= \mathbb{E}[(X_1 + X_2 + W_3 + U_1 + U_2 + w[k-1])^2] \\
&= \mathbb{E}[(X_1 + X_2 + W_3 + U_1 + U_2)^2] + \mathbb{E}[w^2[k-1]] \\
&\geq (\sqrt{\mathbb{E}[(X_1 + X_2 + W_3)^2]} - \sqrt{\mathbb{E}[U_1^2]} - \sqrt{\mathbb{E}[U_2^2]})_+^2 + 1 \\
&= (\sqrt{\mathbb{E}[(X_1 + X_2)^2] + \mathbb{E}[W_3^2]} - \sqrt{\mathbb{E}[U_1^2]} - \sqrt{\mathbb{E}[U_2^2]})_+^2 + 1 \qquad (10.86)
\end{aligned}$$

where the last equality comes form the causality. Here, we can see that $\mathbb{E}[(X_1 + X_2)^2]$ does not depend on the inputs from the power-limited intervals.

• Information-Limited Interval: We will bound the remaining state distortion after the information-limited interval. Define y_1' and y_2' as follows:

$$\begin{aligned}
y_1'[k] &= a^{k-1}w[0] + a^{k-2}w[1] + \cdots + w[k-1] + v_1[k] \\
y_2'[k] &= a^{k-1}w[0] + a^{k-2}w[1] + \cdots + w[k-1] + v_2[k].
\end{aligned}$$

Here, $y_1'[k]$, $y_2'[k]$ can be obtained by removing $u_1[1:k-1]$, $u_2[1:k-1]$ from $y_1[k]$, $y_2[k]$, and $u_1[k]$ and $u_2[k]$ are functions of $y_1[1:k]$ and $y_2[1:k]$ respectively. Therefore, we can see that $y_1[1:k], y_2[1:k]$ are functions of $y_1'[1:k], y_2'[1:k]$. Moreover, W_1, $y_1'[1:k_1-1]$, $y_2'[1:k_1-1]$ are jointly Gaussian.

Let

$$\begin{aligned}
W_1' &:= W_1 - \mathbb{E}[W_1 | y_1'[1:k_1-1], y_2'[1:k_1-1]] \\
W_1'' &:= \mathbb{E}[W_1 | y_1'[1:k_1-1], y_2'[1:k_1-1]].
\end{aligned}$$

Then, W_1', W_1'', W_2 are independent Gaussian random variables. Moreover, W_1', W_2 are independent from $y_1'[1:k_1-1], y_2'[1:k_1-1]$. W_1'' is a function of $y_1'[1:k_1-1], y_2'[1:k_1-1]$.

Now, let's lower bound $\mathbb{E}[(X_1+X_2)^2]$. Since Gaussian maximizes the entropy, we have

$$\begin{aligned}
&\frac{1}{2}\log(2\pi e\mathbb{E}[(X_1+X_2)^2] \\
&\geq h(X_1+X_2) \\
&\geq h(X_1+X_2|y_1'[1:k_1-1], y_2'[1:k_1-1], y_2[k_1:k_2-1]) \\
&= h(W_1'+W_1''+U_{11}+U_{12}+W_2+U_{22}|y_1'[1:k_1-1], y_2'[1:k_1-1], y_2[k_1:k_2-1]) \\
&= h(W_1'+W_2|y_1'[1:k_1-1], y_2'[1:k_1-1], y_2[k_1:k_2-1]).
\end{aligned} \tag{10.87}$$

We will first lower bound the variance of W_1'. Notice that

$$\begin{aligned}
\mathbb{E}[y_1'[k]^2] &= a^{2(k-1)} + a^{2(k-2)} + \cdots + 1 + \sigma_{v1}^2 \\
&= \frac{1-a^{2k}}{1-a^2} + \sigma_{v1}^2
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}[y_2'[k]^2] &= a^{2(k-1)} + a^{2(k-2)} + \cdots + 1 + \sigma_{v1}^2 \\
&= \frac{1-a^{2k}}{1-a^2} + \sigma_{v2}^2.
\end{aligned}$$

Thus, we have

$$\begin{aligned}
&I(W_1; y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
&= h(y_1'[1:k_1-1], y_2'[1:k_1-1]) - h(y_1'[1:k_1-1], y_2'[1:k_1-1]|W_1) \\
&\leq \sum_{1\leq i\leq k_1-1} h(y_1'[i]) + \sum_{1\leq i\leq k_1-1} h(y_2'[i]) - \sum_{1\leq i\leq k_1-1} h(v_1[i]) - \sum_{1\leq i\leq k_1-1} h(v_2[i]) \\
&\leq \sum_{1\leq k\leq k_1-1} \frac{1}{2}\log(\frac{\frac{1-a^{2k}}{1-a^2}+\sigma_{v1}^2}{\sigma_{v1}^2}) + \sum_{1\leq k\leq k_1-1} \frac{1}{2}\log(\frac{\frac{1-a^{2k}}{1-a^2}+\sigma_{v2}^2}{\sigma_{v2}^2}) \\
&= \frac{1}{2}\log(\prod_{1\leq k\leq k_1-1} \frac{\frac{1-a^{2k}}{1-a^2}+\sigma_{v1}^2}{\sigma_{v1}^2}) + \frac{1}{2}\log(\prod_{1\leq k\leq k_1-1} \frac{\frac{1-a^{2k}}{1-a^2}+\sigma_{v2}^2}{\sigma_{v2}^2}) \\
&\overset{(A)}{\leq} \frac{k_1-1}{2}\log(\frac{1}{k_1-1}\sum_{1\leq k\leq k_1-1} \frac{\frac{1-a^{2k}}{1-a^2}+\sigma_{v1}^2}{\sigma_{v1}^2}) + \frac{k_1-1}{2}\log(\frac{1}{k_1-1}\sum_{1\leq k\leq k_1-1} \frac{\frac{1-a^{2k}}{1-a^2}+\sigma_{v2}^2}{\sigma_{v2}^2}) \\
&= \frac{k_1-1}{2}\log(1+\frac{1}{k_1-1}\sum_{1\leq k\leq k_1-1} \frac{\frac{1-a^{2k}}{1-a^2}}{\sigma_{v1}^2}) + \frac{k_1-1}{2}\log(1+\frac{1}{k_1-1}\sum_{1\leq k\leq k_1-1} \frac{\frac{1-a^{2k}}{1-a^2}}{\sigma_{v2}^2}) \\
&\leq \frac{k_1-1}{2}\log(1+\frac{1}{k_1-1}\sum_{1\leq k\leq k_1-1} \frac{\frac{1-a^{2(k_1-1)}}{1-a^2}}{\sigma_{v1}^2}) + \frac{k_1-1}{2}\log(1+\frac{1}{k_1-1}\sum_{1\leq k\leq k_1-1} \frac{\frac{1-a^{2(k_1-1)}}{1-a^2}}{\sigma_{v2}^2}) \\
&= \frac{k_1-1}{2}\log(1+\frac{1}{\sigma_{v1}^2}\frac{1-a^{2(k_1-1)}}{1-a^2}) + \frac{k_1-1}{2}\log(1+\frac{1}{\sigma_{v2}^2}\frac{1-a^{2(k_1-1)}}{1-a^2}).
\end{aligned} \tag{10.88}$$

(A): Arithmetic-Geometric mean

Let's denote the last equation as I. We also have

$$\begin{aligned}\mathbb{E}[W_1^2] &= a^{2(k-1)} + \cdots + a^{2(k-k_1+1)} \\ &= a^{2(k-k_1+1)}(a^{2(k_1-2)} + \cdots + 1) \\ &= a^{2(k-k_1+1)}\frac{1-a^{2(k_1-1)}}{1-a^2}. \end{aligned} \tag{10.89}$$

Now, we can bound the variance of a Gaussian random variable W_1' as follows:

$$\begin{aligned}&\frac{1}{2}\log(2\pi e\mathbb{E}[W_1'^2]) = h(W_1') \\ &\geq h(W_1'|y_1'[1:k_1-1], y_2'[1:k_1-1]) \\ &= h(W_1|y_1'[1:k_1-1], y_2'[1:k_1-1]) \\ &= h(W_1) - I(W_1; y_1'[1:k_1-1], y_2'[1:k_1-1]) \\ &\geq \frac{1}{2}\log(2\pi e a^{2(k-k_1+1)}\frac{1-a^{2(k_1-1)}}{1-a^2}) - I\end{aligned}$$

where the last inequality follows from (10.88) and (10.89).

Thus,

$$\mathbb{E}[W_1'^2] \geq \frac{a^{2(k-k_1+1)}\frac{1-a^{2(k_1-1)}}{1-a^2}}{2^{2I}} \tag{10.90}$$

and denote the last term as Σ. Since W_1' is Gaussian, we can write $W_1' = W_1''' + W_1''''$ where $W_1''' \sim \mathcal{N}(0, \Sigma)$, and W_1''', W_1'''' are independent.

Moreover, we also have

$$\begin{aligned}\mathbb{E}[W_2^2] &= a^{2(k-k_1)} + \cdots + a^{2(k-k_2+1)} \\ &= a^{2(k-k_2+1)}\frac{1-a^{2(k_2-k_1)}}{1-a^2}.\end{aligned} \tag{10.91}$$

By (10.87), we have

$$
\begin{aligned}
&\frac{1}{2}\log(2\pi e\mathbb{E}[(X_1+X_2)^2]) \\
&\geq h(W_1'+W_2|y_1'[1:k_1-1], y_2'[1:k_1-1], y_2[k_1:k_2-1]) \\
&\geq h(W_1'+W_2|W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1], y_2[k_1:k_2-1]) \\
&= h(W_1'''+W_2|W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1], y_2[k_1:k_2-1]) \\
&= h(W_1'''+W_2|W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
&- I(W_1'''+W_2; y_2[k_1:k_2-1]|W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
&= h(W_1'''+W_2) \\
&- I(W_1'''+W_2; y_2[k_1:k_2-1]|W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
&\geq \frac{1}{2}\log(2\pi e(\Sigma + a^{2(k-k_2+1)}\frac{1-a^{2(k_2-k_1)}}{1-a^2})) \\
&- I(W_1'''+W_2; y_2[k_1:k_2-1]|W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \qquad (10.92)
\end{aligned}
$$

where the last inequality comes from the fact that W_1''' and W_2 are independent Gaussian, and (10.90) and (10.91).

Now, the question boils down to the upper bound of the last mutual information term, which can be understood as the information contained in the second controller's observation in Witsenhausen's interval.

• Second controller's observation in Witsenhausen's interval: We will bound the amount of information contained in the second controller's observation in Witsenhausen's interval. For $n \geq k_1$, define

$$
\begin{aligned}
y_2''[n] &:= a^{n-k}W_1''' + a^{n-k_1}w[k_1-1] + a^{n-k_1-1}w[k_1] + \cdots + w[n-1] \\
&\quad + a^{n-k_1-1}u_1[k_1] + \cdots + u_1[n-1] + v_2[n].
\end{aligned}
$$

Notice the relationship between $y_2[n]$ and $y_2''[n]$ is

$$
\begin{aligned}
y_2[n] &= y_2''[n] + a^{n-k_1-1}u_2[k_1] + \cdots + u_2[n-1] \\
&\quad + a^{n-k}W_1'''' + a^{n-k}\mathbb{E}[W_1|y_1'[1:k_1-1], y_2'[1:k_1-1]]. \qquad (10.93)
\end{aligned}
$$

The mutual information in (10.92) is bounded as follows:

$$\begin{aligned}
&I(W_1''' + W_2; y_2[k_1 : k_2-1] | W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
&= h(y_2[k_1:k_2-1]|W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
&- h(y_2[k_1:k_2-1]|W_1'''+W_2, W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
&= \sum_{k_1 \le i \le k_2-1} h(y_2[i]|y_2[k_1:i-1], W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
&- \sum_{k_1 \le i \le k_2-1} h(y_2[i]|y_2[k_1:i-1], W_1'''+W_2, W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
&\overset{(A)}{=} \sum_{k_1 \le i \le k_2-1} h(y_2''[i]|y_2[k_1:i-1], W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
&- \sum_{k_1 \le i \le k_2-1} h(y_2[i]|y_2[k_1:i-1], W_1'''+W_2, W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
&\overset{(B)}{\le} \sum k_1 \le i \le k_2 - 1 h(y_2''[i]) - \sum_{k_1 \le i \le k_2-1} h(v_2[i]) \\
&\le \sum_{k_1 \le i \le k_2-1} \frac{1}{2}\log(2\pi e \mathbb{E}[y_2''[i]^2]) - \sum_{k_1 \le i \le k_2-1} \frac{1}{2}\log(2\pi e \sigma_{v2}^2)
\end{aligned} \tag{10.94}$$

(A): Since $y_2[1:k_1-1]$ is a function of $y_2'[1:k_1-1]$, $u_2[k_1], \cdots, u_2[i]$ are functions of $y_2[k_1 : i-1], y_2'[1:k_1-1]$. Thus, all the terms in (10.93) except $y_2''[i]$ can be vanished by the conditioned.
(B): By causality, $v_2[i]$ is independent from all conditioning random variables.

First, let's bound the variance of $y_2''[n]$. By Lemma 4.1 of Chapter 4, we have

$$\begin{aligned}
\mathbb{E}[y_2''[n]^2] &\le 2\mathbb{E}[(a^{n-k}W_1''' + a^{n-k_1}w[k_1-1] + a^{n-k_1-1}w[k_1] + \cdots + w[n-1])^2] \\
&+ 2\mathbb{E}[(a^{n-k_1-1}u_1[k_1] + \cdots + u_1[n-1])^2] + \sigma_{v2}^2 \\
&= 2(a^{2(n-k)}\Sigma + a^{2(n-k_1)} + \cdots + 1) \\
&+ 2\mathbb{E}[(a^{n-k_1-1}u_1[k_1] + \cdots + u_1[n-1])^2] + \sigma_{v2}^2
\end{aligned}$$

Here, by Lemma 4.1 of Chapter 4, we have

$$\begin{aligned}
&\mathbb{E}[(a^{n-k_1-1}u_1[k_1] + \cdots + u_1[n-1])^2] \\
&\le (\sqrt{a^{2(n-k_1-1)}\mathbb{E}[u_1^2[k_1]]} + \cdots + \sqrt{\mathbb{E}[u_1^2[n-1]]})^2 \\
&\overset{(A)}{\le} (a^{(n-k_1-1)} + a^{(n-k_1-2)} + \cdots + 1)(a^{(n-k_1-1)}\mathbb{E}[u_1^2[k_1]] + a^{(n-k_1-2)}\mathbb{E}[u_1^2[k_1+1]] + \cdots + \mathbb{E}[u_1^2[n-1]]) \\
&= \frac{1-a^{n-k_1}}{1-a} \cdot a^{n-k}(a^{k-k_1-1}\mathbb{E}[u_1^2[k_1]] + a^{k-k_1-2}\mathbb{E}[u_1^2[k_1+1]] + \cdots + a^{k-n}\mathbb{E}[u_1^2[n-1]]) \\
&\le \frac{1-a^{n-k_1}}{1-a} \cdot a^{n-k}\frac{1-a^{k-k_1}}{1-a}\widetilde{P_1} \\
&= a^{n-k}\frac{(1-a^{n-k})(1-a^{k-k_1})}{(1-a)^2}\widetilde{P_1}.
\end{aligned}$$

(A): Cauchy-Schwarz inequality

Thus, the variance of $y_2''[n]$ is bounded as:

$$\mathbb{E}[y_2''[n]^2] \leq 2a^{2(n-k)}\Sigma + 2\frac{1-a^{2(n-k_1+1)}}{1-a^2} + 2a^{n-k}\frac{(1-a^{n-k})(1-a^{k-k_1})}{(1-a)^2}\widetilde{P_1} + \sigma_{v2}^2.$$

Therefore, we have

$$\begin{aligned}
&\sum_{k_1\leq n\leq k_2-1} \mathbb{E}[y_2''[n]^2] \\
&\leq \sum_{k_1\leq n\leq k_2-1} 2a^{2(n-k)}\Sigma + 2\frac{1-a^{2(n-k_1+1)}}{1-a^2} + 2a^{n-k}\frac{(1-a^{n-k})(1-a^{k-k_1})}{(1-a)^2}\widetilde{P_1} + \sigma_{v2}^2 \\
&\leq 2(a^{2(k_1-k)} + \cdots + a^{2(k_2-1-k)})\Sigma + \sum_{k_1\leq n\leq k_2-1} 2\frac{1-a^{2(k_2-1-k_1+1)}}{1-a^2} \\
&+ \sum_{k_1\leq n\leq k_2-1} 2a^{n-k}\frac{(1-a^{k_2-1-k_1})(1-a^{k-k_1})}{(1-a)^2}\widetilde{P_1} + (k_2-k_1)\sigma_{v2}^2 \\
&\leq 2a^{2(k_1-k)}\frac{1-a^{2(k_2-k_1)}}{1-a^2}\Sigma + 2(k_2-k_1)\frac{1-a^{2(k_2-1-k_1+1)}}{1-a^2} \\
&+ 2a^{k_1-k}\frac{1-a^{k_2-k_1}}{1-a}\frac{(1-a^{k_2-1-k_1})(1-a^{k-k_1})}{(1-a)^2}\widetilde{P_1} + (k_2-k_1)\sigma_{v2}^2. \qquad (10.95)
\end{aligned}$$

Therefore, by (10.94) and (10.95) we conclude

$$\begin{aligned}
&I(W_1''' + W_2; y_2[k_1:k_2-1]|W'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
&\leq \sum_{k_1\leq n\leq k_2-1} \frac{1}{2}\log(\frac{\mathbb{E}[y_2''[n]^2]}{\sigma_{v2}^2}) \\
&= \frac{1}{2}\log(\prod_{k_1\leq n\leq k_2-1} \frac{\mathbb{E}[y_2''[n]^2]}{\sigma_{v2}^2}) \\
&\overset{(A)}{\leq} \frac{k_2-k_1}{2}\log(\frac{1}{k_2-k_1}\sum_{k_1\leq n\leq k_2-1}\frac{\mathbb{E}[y_2''[n]^2]}{\sigma_{v2}^2}) \\
&\leq \frac{k_2-k_1}{2}\log(1+\frac{1}{(k_2-k_1)\sigma_{v2}^2}(2a^{2(k_1-k)}\frac{1-a^{2(k_2-k_1)}}{1-a^2}\Sigma + 2(k_2-k_1)\frac{1-a^{2(k_2-1-k_1+1)}}{1-a^2} \\
&+ 2a^{k_1-k}\frac{1-a^{k_2-k_1}}{1-a}\frac{(1-a^{k_2-1-k_1})(1-a^{k-k_1})}{(1-a)^2}\widetilde{P_1}))
\end{aligned}$$

(A): Arithmetic-Geometric mean

Denote the last equation as $I'(\widetilde{P_1})$. By (10.92) we conclude

$$\begin{aligned}
&\frac{1}{2}\log(2\pi e\mathbb{E}[(X_1+X_2)^2] \\
&\geq h(W_1''' + W_2) - I(W_1''' + W_2; y_2[k_1:k_2-1]|W_1'''', y_1'[1:k_1-1], y_2'[1:k_1-1]) \\
&\geq \frac{1}{2}\log(2\pi e(\Sigma + a^{2(k-k_2+1)}\frac{1-a^{2(k_2-k_1)}}{1-a^2})) - I'(\widetilde{P_1})
\end{aligned}$$

which implies

$$\mathbb{E}[(X_1+X_2)^2] \geq \frac{\Sigma + a^{2(k-k_2+1)}\frac{1-a^{2(k_2-k_1)}}{1-a^2}}{2^{2I'(\widetilde{P_1})}}. \tag{10.96}$$

• Final lower bound: Now, we can merge the inequalities to prove the lemma. The variance of W_3 is given as follows:

$$\begin{aligned}\mathbb{E}[W_3^2] &= a^{2(k-k_2)} + \cdots + a^2 \\ &= a^2\frac{1-a^{2(k-k_2)}}{1-a^2}.\end{aligned} \tag{10.97}$$

By Lemma 4.1 of Chapter 4, the variance of U_1 is bounded as follows:

$$\begin{aligned}\mathbb{E}[U_1^2] &\leq (\sqrt{a^{2(k-k_1-1)}\mathbb{E}[u_1^2[k_1]]} + \cdots + \sqrt{\mathbb{E}[u_1^2[k-1]]})^2 \\ &\overset{(A)}{\leq} (a^{(k-k_1-1)} + a^{(k-k_1-2)} + \cdots + 1)(a^{(k-k_1-1)}\mathbb{E}[u_1^2[k_1]] + a^{(k-k_1-2)}\mathbb{E}[u_1^2[k_1+1]] + \cdots + \mathbb{E}[u_1^2[k-1]]) \\ &= \frac{1-a^{k-k_1}}{1-a}(a^{(k-k_1-1)}\mathbb{E}[u_1^2[k_1]] + a^{(k-k_1-2)}\mathbb{E}[u_1^2[k_1+1]] + \cdots + \mathbb{E}[u_1^2[k-1]]) \\ &= \frac{1-a^{k-k_1}}{1-a}\frac{1-a^{k-k_1}}{1-a}\widetilde{P_1} \\ &= (\frac{1-a^{k-k_1}}{1-a})^2\widetilde{P_1}.\end{aligned} \tag{10.98}$$

(A): Cauchy-Schwarz inequality

Likewise, the variance of U_2 can be bounded as

$$\mathbb{E}[U_2^2] \leq (\frac{1-a^{k-k_2}}{1-a})^2\widetilde{P_2} \tag{10.99}$$

By plugging in (10.96), (10.97), (10.98), (10.99) into (10.86), we finally prove the lemma. □

Proof of Corollary 5.6 of Page 252. For simplicity, we will prove for $0.9 \leq a < 1$. The proof for $-1 < a \leq -0.9$ is simply follows by replacing a by $|a|$.

First, we can notice

$$\begin{aligned}\Sigma_1 &= \frac{(a^2-1)\sigma_{v1}^2 - 1 + \sqrt{((a^2-1)\sigma_{v1}^2-1)^2 + 4a^2\sigma_{v1}^2}}{2a^2} \\ &= \frac{4a^2\sigma_{v1}^2}{2a^2(-(a^2-1)\sigma_{v1}^2 + 1 + \sqrt{((a^2-1)\sigma_{v1}^2-1)^2 + 4a^2\sigma_{v1}^2})} \\ &= \frac{2\sigma_{v1}^2}{(1-a^2)\sigma_{v1}^2 + 1 + \sqrt{((1-a^2)\sigma_{v1}^2+1)^2 + 4a^2\sigma_{v1}^2}}\end{aligned}$$

Since $0.9 \leq a < 1$, $(1-a^2)\sigma_{v1}^2 \geq 0$. Thus, Σ_1 is upper bounded by

$$\Sigma_1 \leq \frac{2\sigma_{v1}^2}{\sqrt{4a^2\sigma_{v1}^2}} = \frac{\sigma_{v1}^2}{a\sigma_{v1}} = \frac{\sigma_{v1}}{a} \leq \frac{10}{9}\sigma_{v1}$$

and

$$\Sigma_1 \leq \frac{2\sigma_{v1}^2}{(1-a^2)\sigma_{v1}^2 + (1-a^2)\sigma_{v1}^2}$$
$$= \frac{1}{1-a^2}$$

Likewise, we also have

$$\Sigma_2 \leq \frac{10}{9}\sigma_{v2} \tag{10.100}$$

and

$$\Sigma_2 \leq \frac{1}{1-a^2} \tag{10.101}$$

Proof of (a):

Notice that by $\Sigma_2 \geq 40$, $0.9 \leq a < 1$ and (10.101) we have

$$\frac{\Sigma_2}{40} \leq \frac{1}{40}\frac{1}{1-a^2} < \frac{a^2}{1-a^2}$$
$$\frac{\Sigma_2}{40} \geq 1 \geq a^2 = \frac{a^2-a^4}{1-a^2}$$

Thus, we can find $k \geq 3$ such that

$$\frac{a^2-a^{2(k-1)}}{1-a^2} \leq \frac{\Sigma_2}{40} < \frac{a^2-a^{2k}}{1-a^2} \tag{10.102}$$

Let's put such k and $k_1 = 1, k_2 = k$ as the parameters of Lemma 5.5. Then, the lower bound of Lemma 5.5 reduces to

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq (\sqrt{\frac{\frac{a^2-a^{2k}}{1-a^2}}{2^{2I'(\widetilde{P_1})}}} - \sqrt{(\frac{1-a^{k-1}}{1-a})^2\widetilde{P_1}})_+^2 + 1 \tag{10.103}$$

where

$$I'(\widetilde{P_1}) = \frac{1}{2}\log(1 + \frac{1}{(k-1)\sigma_{v2}^2}(2(k-1)\frac{1-a^{2(k-1)}}{1-a^2} + 2a^{1-k}\frac{1-a^{k-1}}{1-a}\frac{(1-a^{k-2})(1-a^{k-1})}{(1-a)^2}\widetilde{P_1}))^{k-1}.$$

Let's first upper bound $I'(\widetilde{P_1})$. By (10.101) and (10.102), we first have

$$\frac{a^2-a^{2(k-1)}}{1-a^2} \leq \frac{\Sigma_2}{40} \leq \frac{1}{40}\frac{1}{1-a^2}$$
$$(\Rightarrow) a^2 - a^{2(k-1)} \leq \frac{1}{40}$$
$$(\Leftrightarrow) a^2 - \frac{1}{40} \leq a^{2(k-1)}$$
$$(\Rightarrow) (\frac{9}{10})^2 - \frac{1}{40} \leq a^{2(k-1)} (\because 0.9 \leq a < 1)$$
$$(\Leftrightarrow) \frac{157}{200} \leq a^{2(k-1)}. \tag{10.104}$$

We also have

$$a^{2(k-1)}(k-1) \leq 1 + a^2 + a^4 + \cdots + a^{2(k-2)} = \frac{1 - a^{2(k-1)}}{1 - a^2} \tag{10.105}$$

where the first inequality comes from that $0.9 \leq a < 1$ so $a^{2(k-1)} \leq 1, \cdots, a^{2(k-1)} \leq a^{2(k-2)}$.

Therefore, by (10.104) and (10.105)

$$\begin{aligned} &\frac{157}{200}(k-1) \leq \frac{1 - a^{2(k-1)}}{1 - a^2} \\ &(\Rightarrow) k - 1 \leq \frac{200}{157}\frac{1 - a^{2(k-1)}}{1 - a^2} \end{aligned} \tag{10.106}$$

Moreover, we also have

$$\begin{aligned} \frac{1 - a^{2(k-1)}}{a^2 - a^{2(k-1)}} &= \frac{1 - a^2}{a^2 - a^{2(k-1)}} + \frac{a^2 - a^{2(k-1)}}{a^2 - a^{2(k-1)}} \\ &= \frac{1 - a^2}{a^2 - a^{2(k-1)}} + 1 \\ &\leq \frac{1 - a^2}{a^2 - a^4} + 1 (\because k \geq 3) \\ &= \frac{1}{a^2} + 1 \\ &\leq (\frac{10}{9})^2 + 1 = \frac{181}{81} (\because 0.9 \leq a < 1) \end{aligned}$$

which implies

$$\frac{1 - a^{2(k-1)}}{1 - a^2} \leq \frac{181}{81}\frac{a^2 - a^{2(k-1)}}{1 - a^2}. \tag{10.107}$$

We also have

$$\begin{aligned} \frac{1 - a^{k-1}}{1 - a} &\leq \frac{1 - a^{k-1}}{1 - a}\frac{2}{1 + a} (\because 0.9 \leq a < 1) \\ &\leq \frac{1 - a^{2(k-2)}}{1 - a}\frac{2}{1 + a} (\because k \geq 3 \text{ implies } 2(k-2) \geq k - 1) \\ &\leq \frac{1 - a^{2(k-2)}}{1 - a}\frac{2}{1 + a}\frac{a^2}{0.9^2} (\because 0.9 \leq a < 1) \\ &= \frac{2}{0.9^2}\frac{a^2 - a^{2(k-1)}}{1 - a^2}. \end{aligned} \tag{10.108}$$

Therefore, the terms in $I'(\widetilde{P_1})$ are upper bounded as:

$$\begin{aligned} 2(k-1)\frac{1 - a^{2(k-1)}}{1 - a^2} &\leq 2\frac{200}{157}(\frac{1 - a^{2(k-1)}}{1 - a^2})^2 (\because (10.106)) \\ &\leq 2\frac{200}{157}(\frac{181}{81}\frac{a^2 - a^{2(k-1)}}{1 - a^2})^2 (\because (10.107)) \\ &\leq 2\frac{200}{157}(\frac{181}{81}\frac{\Sigma_2}{40})^2 (\because (10.102)) \end{aligned} \tag{10.109}$$

and

$$
\begin{aligned}
&2a^{1-k}\frac{1-a^{k-1}}{1-a}\frac{(1-a^{k-2})(1-a^{k-1})}{(1-a)^2}\widetilde{P_1} \\
&\leq 2\sqrt{\frac{200}{157}}(\frac{1-a^{k-1}}{1-a})^3\widetilde{P_1}(\because (10.104) \text{ and } a^{k-1}\leq a^{k-2}) \\
&\leq 2\sqrt{\frac{200}{157}}(\frac{2}{0.9^2}(\frac{a^2-a^{2(k-1)}}{1-a^2}))^3\widetilde{P_1}(\because (10.108)) \\
&\leq 2\sqrt{\frac{200}{157}}(\frac{2}{0.9^2}\frac{\Sigma_2}{40})^3\frac{1}{\Sigma_2}(\because (10.102) \text{ and the assumption } \widetilde{P_1}\leq\frac{1}{\Sigma_2}) \\
&= 2\sqrt{\frac{200}{157}}(\frac{5}{81})^3\Sigma_2^2 \qquad (10.110)
\end{aligned}
$$

Now, we can upper bound $I'(\widetilde{P_1})$ by

$$
\begin{aligned}
I'(\widetilde{P_1}) &\overset{(A)}{\leq} \frac{1}{2}\log(1+\frac{\Sigma_2^2}{(k-1)\sigma_{v2}^2}(2\frac{200}{157}(\frac{181}{81\cdot 40})^2+2\sqrt{\frac{200}{157}}(\frac{5}{81})^3))^{k-1} \\
&\overset{(B)}{\leq} \frac{1}{2}\log(1+\frac{1}{(k-1)}(\frac{10}{9})^2(2\frac{200}{157}(\frac{181}{81\cdot 40})^2+2\sqrt{\frac{200}{157}}(\frac{5}{81})^3))^{k-1} \\
&\leq \frac{1}{2}\log(1+\frac{0.010471667...}{k-1})^{k-1} \\
&\leq \frac{1}{2}\log e^{0.01047} \qquad (10.111)
\end{aligned}
$$

(A): (10.109), (10.110)

(B): (10.100)

Moreover, we also have

$$
\begin{aligned}
(\frac{1-a^{k-1}}{1-a})^2\widetilde{P_1} &\leq (\frac{2}{0.9^2}\frac{a^2-a^{2(k-1)}}{1-a^2})^2\widetilde{P_1}(\because (10.108)) \\
&\leq (\frac{2}{0.9^2}\frac{\Sigma_2}{40})^2\frac{1}{\Sigma_2}(\because (10.102) \text{ and the assumption } \widetilde{P_1}\leq\frac{1}{\Sigma_2}) \\
&= (\frac{5}{81})^2\Sigma_2. \qquad (10.112)
\end{aligned}
$$

Finally, by plugging (10.111), (10.112) into (10.103) we can conclude

$$
\begin{aligned}
D_L(\widetilde{P_1},\widetilde{P_2}) &\geq (\sqrt{\frac{\Sigma_2}{40e^{0.01047}}}-\sqrt{(\frac{5}{81})^2\Sigma_2})_+^2+1 \\
&\geq 0.0091316992...\Sigma_2+1.
\end{aligned}
$$

Proof of (b):

Notice that by $\Sigma_2\geq 40$, $\frac{1}{\Sigma_2}\leq\widetilde{P_1}\leq\frac{1}{40}$, (10.101),

$$
\begin{aligned}
\frac{1}{40\widetilde{P_1}} &\leq \frac{\Sigma_2}{40}\leq\frac{1}{40}\frac{1}{1-a^2}<\frac{a^2}{1-a^2} \\
\frac{1}{40\widetilde{P_1}} &\geq 1\geq a^2=\frac{a^2-a^4}{1-a^2}
\end{aligned}
$$

Thus, we can find $k \geq 3$ such that

$$\frac{a^2 - a^{2(k-1)}}{1 - a^2} \leq \frac{1}{40\widetilde{P_1}} < \frac{a^2 - a^{2k}}{1 - a^2} \tag{10.113}$$

Let's put such k and $k_1 = 1$, $k_2 = k$ as the parameters of Lemma 5.5. Then, the lower bound of Lemma 5.5 reduces to

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq (\sqrt{\frac{\frac{a^2 - a^{2k}}{1-a^2}}{2^{2I'(\widetilde{P_1})}}} - \sqrt{(\frac{1 - a^{k-1}}{1 - a})^2 \widetilde{P_1}})_+^2 + 1$$

First, we will upper bound $I'(\widetilde{P_1})$. Since $\frac{1}{40\widetilde{P_1}} \leq \frac{\Sigma_2}{40}$, we still have (10.104), (10.105), (10.106) which are

$$\frac{157}{200} \leq a^{2(k-1)} \tag{10.114}$$

$$k - 1 \leq \frac{200}{157} \frac{1 - a^{2(k-1)}}{1 - a^2}. \tag{10.115}$$

Since $k \geq 3$, we still have (10.107) and (10.108) which are

$$\frac{1 - a^{2(k-1)}}{1 - a^2} \leq \frac{181}{81} \frac{a^2 - a^{2(k-1)}}{1 - a^2} \tag{10.116}$$

$$\frac{1 - a^{k-1}}{1 - a} \leq \frac{2}{0.9^2} \frac{a^2 - a^{2(k-1)}}{1 - a^2}. \tag{10.117}$$

Then, the terms in $I'(\widetilde{P_1})$ are upper bounded as:

$$\begin{aligned} 2(k-1)\frac{1 - a^{2(k-1)}}{1 - a^2} &\leq 2\frac{200}{157}(\frac{1 - a^{2(k-1)}}{1 - a^2})^2 (\because (10.115)) \\ &\leq 2\frac{200}{157}(\frac{181}{81}\frac{a^2 - a^{2(k-1)}}{1 - a^2})^2 (\because (10.116)) \\ &\leq 2\frac{200}{157}(\frac{181}{81}\frac{1}{40\widetilde{P_1}})^2 (\because (10.113)) \\ &\leq 2\frac{200}{157}(\frac{181}{81}\frac{1}{40})^2 \Sigma_2^2 (\because \text{Assumption } \frac{1}{\widetilde{P_1}} \leq \Sigma_2) \end{aligned} \tag{10.118}$$

and

$$\begin{aligned} &2a^{1-k}\frac{1 - a^{k-1}}{1 - a}\frac{(1 - a^{k-2})(1 - a^{k-1})}{(1 - a)^2}\widetilde{P_1} \\ &\leq 2\sqrt{\frac{200}{157}}(\frac{1 - a^{k-1}}{1 - a})^3 \widetilde{P_1} (\because (10.114)) \\ &\leq 2\sqrt{\frac{200}{157}}(\frac{2}{0.9^2}(\frac{a^2 - a^{2(k-1)}}{1 - a^2}))^3 \widetilde{P_1} (\because (10.117)) \\ &\leq 2\sqrt{\frac{200}{157}}(\frac{2}{0.9^2}\frac{1}{40\widetilde{P_1}})^3 \widetilde{P_1} (\because (10.113)) \\ &= 2\sqrt{\frac{200}{157}}(\frac{5}{81})^3 \frac{1}{\widetilde{P_1}^2} \\ &\leq 2\sqrt{\frac{200}{157}}(\frac{5}{81})^3 \Sigma_2^2 . (\because \text{Assumption } \frac{1}{\widetilde{P_1}} \leq \Sigma_2) \end{aligned} \tag{10.119}$$

Therefore, by (10.118) and (10.119), $I'(\widetilde{P_1})$ is upper bounded as:

$$\begin{aligned}
I'(\widetilde{P_1}) &\leq \frac{1}{2}\log(1+\frac{\Sigma_2^2}{(k-1)\sigma_{v2}^2}(2\frac{200}{157}(\frac{181}{81\cdot 40})^2+2\sqrt{\frac{200}{157}}(\frac{5}{81})^3))^{k-1} \\
&\leq \frac{1}{2}\log(1+\frac{1}{(k-1)}(\frac{10}{9})^2(2\frac{200}{157}(\frac{181}{81\cdot 40})^2+2\sqrt{\frac{200}{157}}(\frac{5}{81})^3))^{k-1}(\because (10.100)) \\
&\leq \frac{1}{2}\log(1+\frac{0.010471667...}{k-1})^{k-1} \\
&\leq \frac{1}{2}\log e^{0.01047}. \qquad (10.120)
\end{aligned}$$

Moreover, we also have

$$\begin{aligned}
(\frac{1-a^{k-1}}{1-a})^2\widetilde{P_1} &\leq (\frac{2}{0.9^2}\frac{a^2-a^{2(k-1)}}{1-a^2})^2\widetilde{P_1}(\because (10.117)) \\
&\leq (\frac{2}{0.9^2}\frac{1}{40\widetilde{P_1}})^2\widetilde{P_1}(\because (10.113)) \\
&= (\frac{5}{81})^2\frac{1}{\widetilde{P_1}}. \qquad (10.121)
\end{aligned}$$

Finally, by (10.113), (10.120), (10.121), we can conclude

$$\begin{aligned}
D_L(\widetilde{P_1},\widetilde{P_2}) &\geq (\sqrt{\frac{1}{40e^{0.01047}\widetilde{P_1}}}-\sqrt{(\frac{5}{81})^2\frac{1}{\widetilde{P_1}}})_+^2+1 \\
&\geq \frac{0.0091316992...}{\widetilde{P_1}}+1
\end{aligned}$$

Proof of (c):

Let $P := \max(\widetilde{P_1},\widetilde{P_2})$. Notice that since $\frac{1-a^2}{20} \leq P \leq \frac{1}{40}$ and $0.9 \leq a < 1$ we have

$$\begin{aligned}
\frac{1}{40P} &\leq \frac{1}{2(1-a^2)} < \frac{a^2}{1-a^2} \\
\frac{1}{40P} &\geq 1 \geq a^2 = \frac{a^2-a^4}{1-a^2}
\end{aligned}$$

Therefore, we can find $k \geq 3$ such that

$$\frac{a^2-a^{2(k-1)}}{1-a^2} \leq \frac{1}{40P} < \frac{a^2-a^{2k}}{1-a^2} \qquad (10.122)$$

Let's put such k and $k_1 = k_2 = 1$ as the parameters of Lemma 5.5. Then, the lower bound of Lemma 5.5 reduces to

$$D_L(\widetilde{P_1},\widetilde{P_2}) \geq (\sqrt{\frac{a^2-a^{2k}}{1-a^2}}-\sqrt{(\frac{1-a^{k-1}}{1-a})^2\widetilde{P_1}}-\sqrt{(\frac{1-a^{k-1}}{1-a})^2\widetilde{P_2}})_+^2+1. \qquad (10.123)$$

Since $k \geq 3$, we still have (10.108) which tells $\frac{1-a^{k-1}}{1-a} \leq \frac{2}{0.9^2}\frac{a^2-a^{2(k-1)}}{1-a^2}$. Thus, by (10.108), we have

$$\begin{aligned}
\frac{1-a^{k-1}}{1-a} &\leq \frac{2}{0.9^2}\frac{a^2-a^{2(k-1)}}{1-a^2} \\
&\leq \frac{2}{0.9^2}\frac{1}{40P}(\because (10.122)) \\
&= \frac{5}{81P} \qquad (10.124)
\end{aligned}$$

Thus, by plugging (10.122), (10.124) into (10.123), we can conclude

$$\begin{aligned}
D_L(\widetilde{P_1}, \widetilde{P_2}) &\geq (\sqrt{\frac{1}{40P}} - \sqrt{(\frac{5}{81P})^2 P} - \sqrt{(\frac{5}{81P})^2 P})_+^2 + 1 \\
&= (\sqrt{\frac{1}{40}} - \sqrt{(\frac{5}{81})^2} - \sqrt{(\frac{5}{81})^2})^2 \frac{1}{P} + 1 \\
&= 0.0012011... \frac{1}{P} + 1
\end{aligned}$$

Proof (d):

Since $\frac{1}{2} < a^4$, there exists $k \geq 3$ such that

$$a^{2k} \leq \frac{1}{2} < a^{2(k-1)} \tag{10.125}$$

Let's put such k and $k_1 = k_2 = 1$ as the parameters of Lemma 5.5. Then, the lower bound of Lemma 5.5 reduces to

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq (\sqrt{\frac{a^2 - a^{2k}}{1 - a^2}} - \sqrt{(\frac{1 - a^{k-1}}{1 - a})^2 \widetilde{P_1}} - \sqrt{(\frac{1 - a^{k-1}}{1 - a})^2 \widetilde{P_2}})_+^2 + 1$$

Here, we have

$$\begin{aligned}
\frac{a^2 - a^{2k}}{1 - a^2} &\geq \frac{a^2 - \frac{1}{2}}{1 - a^2} (\because (10.125)) \\
&\geq \frac{0.9^2 - \frac{1}{2}}{1 - a^2} (\because 0.9 \leq a < 1) \\
&= \frac{0.31}{1 - a^2}
\end{aligned} \tag{10.126}$$

and

$$\begin{aligned}
\frac{1 - a^{k-1}}{1 - a} &\leq \frac{1 - \frac{1}{\sqrt{2}}}{1 - a} (\because (10.125)) \\
&\leq \frac{1 - \frac{1}{\sqrt{2}}}{1 - a} \frac{2}{1 + a} (\because 0.9 \leq a < 1) \\
&= \frac{2(1 - \frac{1}{\sqrt{2}})}{1 - a^2}
\end{aligned} \tag{10.127}$$

Finally, by the assumption $\max(\widetilde{P_1}, \widetilde{P_2}) \leq \frac{1-a^2}{20}$ and (10.126), (10.127) we can conclude

$$\begin{aligned}
D_L(\widetilde{P_1}, \widetilde{P_2}) &\geq (\sqrt{\frac{0.31}{1 - a^2}} - \sqrt{(\frac{2(1 - \frac{1}{\sqrt{2}})}{1 - a^2})^2 \frac{1 - a^2}{20}} - \sqrt{(\frac{2(1 - \frac{1}{\sqrt{2}})}{1 - a^2})^2 \frac{1 - a^2}{20}})_+^2 + 1 \\
&= (\sqrt{0.31} - \sqrt{(2(1 - \frac{1}{\sqrt{2}}))^2 \frac{1}{20}} - \sqrt{(2(1 - \frac{1}{\sqrt{2}}))^2 \frac{1}{20}})^2 \frac{1}{1 - a^2} + 1 \\
&= \frac{0.0869099...}{1 - a^2} + 1
\end{aligned}$$

Proof of (e):

As mentioned in the proof of Corollary 5.4 (j), the centralized controller's distortion that has both observations $y_1[n], y_2[n]$ and has no input power constraints is a lower bound on the decentralized controller's distortion.

Let $y_1'[n] := x[n] + v_1'[n]$ and $y_2'[n] := x[n] + v_2'[n]$ where $v_1'[n] \sim \mathcal{N}(0, \sigma_1^2)$ and $v_2'[n] \sim \mathcal{N}(0, \sigma_1^2)$ are i.i.d. random variables. Just like the proof of Corollary 5.4 (j), the performance of the centralized controller with both observations is equivalent to a centralized controller with observation $\frac{y_1'[n]+y_2'[n]}{2}$ by the maximum ratio combining.

Let Σ_E be the estimation error of the Kalman filtering with a scalar observation $\frac{y_1'[n]+y_2'[n]}{2}$. By Lemma 5.1,

$$\begin{aligned}\Sigma_E &= \frac{(a^2-1)\frac{\sigma_{v1}^2}{2} - 1 + \sqrt{((a^2-1)\frac{\sigma_{v1}^2}{2} - 1)^2 + 4a^2\frac{\sigma_{v1}^2}{2}}}{2a^2} \\ &= \frac{4a^2\frac{\sigma_{v1}^2}{2}}{2a^2(-(a^2-1)\frac{\sigma_{v1}^2}{2} + 1 + \sqrt{((a^2-1)\frac{\sigma_{v1}^2}{2} - 1)^2 + 4a^2\frac{\sigma_{v1}^2}{2}})} \\ &= \frac{\sigma_{v1}^2}{(1-a^2)\frac{\sigma_{v1}^2}{2} + 1 + \sqrt{((1-a^2)\frac{\sigma_{v1}^2}{2} + 1)^2 + 4a^2\frac{\sigma_{v1}^2}{2}}}.\end{aligned}$$

Here, we have

$$\Sigma_E \leq \frac{\sigma_{v1}^2}{(1-a^2)\frac{\sigma_{v1}^2}{2} + (1-a^2)\frac{\sigma_{v1}^2}{2}} = \frac{1}{1-a^2}. \tag{10.128}$$

Then, for all $\widetilde{P_1}$ and $\widetilde{P_2}$, the cost of the decentralized controllers is lower bounded as follows:

$$\begin{aligned}D_L(\widetilde{P_1}, \widetilde{P_2}) &\overset{(A)}{\geq} \inf_{k:|a-k|<1} \frac{(2ak-k^2)\Sigma_E + 1}{1-(a-k)^2} \\ &= \inf_{k:|a-k|<1} \frac{a^2-1}{1-(a-k)^2}\Sigma_E + \frac{1-a^2+2ak-k^2}{1-(a-k)^2}\Sigma_E + \frac{1}{1-(a-k)^2} \\ &= \inf_{k:|a-k|<1} \frac{a^2-1}{1-(a-k)^2}\Sigma_E + \Sigma_E + \frac{1}{1-(a-k)^2} \end{aligned} \tag{10.129}$$

(A): The decentralized control cost is larger than the centralized controller's cost with the observation $\frac{y_1'[n]+y_2'[n]}{2}$. Moreover, when $|a-k| \geq 1$ the centralized control system is unstable, and the cost diverges to infinity. When $|a-k| < 1$, the cost analysis follows from Lemma 5.1.

Let $k^\star$ be k achieving the infimum of (10.129). Then, for all $\widetilde{P_1}, \widetilde{P_2} \geq 0$ we have

$$\begin{aligned}D_L(\widetilde{P_1}, \widetilde{P_2}) &\geq \frac{a^2-1}{1-(a-k^\star)^2}\Sigma_E + \Sigma_E + \frac{1}{1-(a-k^\star)^2} \\ &\geq \frac{a^2-1}{1-(a-k^\star)^2}\frac{1}{1-a^2} + \Sigma_E + \frac{1}{1-(a-k^\star)^2} (\because (10.128)) \\ &= \Sigma_E. \end{aligned} \tag{10.130}$$

Therefore, Σ_E is a lower bound on $D_L(\widetilde{P_1}, \widetilde{P_2})$, and we will compare this with Σ_1 which is

$$\Sigma_1 = \frac{2\sigma_{v1}^2}{(1-a^2)\sigma_{v1}^2 + 1 + \sqrt{((1-a^2)\sigma_{v1}^2 + 1)^2 + 4a^2\sigma_{v1}^2}}.$$

To this end, let's divide the case based on three quantities $(1-a^2)\frac{\sigma_{v1}^2}{2}, 1, a\frac{\sigma_{v1}}{\sqrt{2}}$.
(i) When $\max((1-a^2)\frac{\sigma_{v1}^2}{2}, 1, a\frac{\sigma_{v1}}{\sqrt{2}}) = (1-a^2)\frac{\sigma_{v1}^2}{2}$,
In this case, by the definition of Σ_E we have

$$\begin{aligned}\Sigma_E &\geq \frac{\sigma_{v1}^2}{2(1-a^2)\frac{\sigma_{v1}^2}{2} + \sqrt{(2(1-a^2)\frac{\sigma_{v1}^2}{2})^2 + 4((1-a^2)\frac{\sigma_{v1}^2}{2})^2}} \\ &= \frac{1}{(1-a^2) + \sqrt{(1-a^2)^2 + (1-a^2)^2}} \\ &= \frac{1}{1+\sqrt{2}}\frac{1}{1-a^2}.\end{aligned} \tag{10.131}$$

By the definition of Σ_1, we also have

$$\begin{aligned}\Sigma_1 &\leq \frac{2\sigma_{v1}^2}{(1-a^2)\sigma_{v1}^2 + (1-a^2)\sigma_{v1}^2} \\ &= \frac{1}{1-a^2}.\end{aligned} \tag{10.132}$$

Therefore, we have

$$\begin{aligned}D_L(\widetilde{P_1}, \widetilde{P_2}) &\geq \Sigma_E (\because (10.130)) \\ &\geq \frac{1}{1+\sqrt{2}}\frac{1}{1-a^2} (\because (10.131)) \\ &\geq \frac{1}{1+\sqrt{2}}\Sigma_1 . (\because (10.132))\end{aligned}$$

(ii) When $\max((1-a^2)\frac{\sigma_{v1}^2}{2}, 1, a\frac{\sigma_{v1}}{\sqrt{2}}) = a\frac{\sigma_{v1}}{\sqrt{2}}$,
In this case, by the definition of Σ_E we have

$$\begin{aligned}\Sigma_E &\geq \frac{\sigma_{v1}^2}{a\frac{\sigma_{v1}}{\sqrt{2}} + a\frac{\sigma_{v1}}{\sqrt{2}} + \sqrt{(a\frac{\sigma_{v1}}{\sqrt{2}} + a\frac{\sigma_{v1}}{\sqrt{2}})^2 + 4a^2\frac{\sigma_{v1}^2}{2}}} \\ &= \frac{\sigma_{v1}}{\frac{2a}{\sqrt{2}} + \sqrt{(\frac{2a}{\sqrt{2}})^2 + 2a^2}} \\ &\geq \frac{\sigma_{v1}}{\frac{2}{\sqrt{2}} + \sqrt{2+2}} (\because 0.9 \leq a < 1) \\ &= \frac{\sigma_{v1}}{\frac{2}{\sqrt{2}} + 2}.\end{aligned} \tag{10.133}$$

By the definition of Σ_1, we also have

$$\begin{aligned}\Sigma_1 &\leq \frac{2\sigma_{v1}^2}{\sqrt{4a^2\sigma_{v1}^2}} = \frac{\sigma_{v1}^2}{a\sigma_{v1}} = \frac{\sigma_{v1}}{a} \\ &\leq \frac{10}{9}\sigma_{v1} . (\because 0.9 \leq a < 1)\end{aligned} \tag{10.134}$$

Therefore, we have

$$\begin{aligned} D_L(P_1,P_2) &\geq \Sigma_E (\because (10.130)) \\ &\geq \frac{\sigma_{v1}}{\frac{2}{\sqrt{2}}+2}(10.133) \\ &\geq \frac{1}{\frac{2}{\sqrt{2}}+2}\frac{9}{10}\Sigma_1.(10.134) \end{aligned}$$

(iii) When $\max((1-a^2)\frac{\sigma_{v1}^2}{2}, 1, a\frac{\sigma_{v1}}{\sqrt{2}}) = 1$,

By the assumption, we have $a\frac{\sigma_{v1}}{\sqrt{2}} \leq 1$. Since $0.9 \leq a < 1$, we can see

$$\sigma_{v1} \leq \frac{\sqrt{2}}{a} \leq \frac{10\sqrt{2}}{9}. \tag{10.135}$$

Furthermore, by the definition of Σ_1, we can see that (10.134) still holds. Therefore, by (10.135), we have

$$\Sigma_1 \leq \frac{10}{9}\sigma_{v1} \leq \frac{10}{9}(\frac{10\sqrt{2}}{9}) \tag{10.136}$$

Moreover, by Lemma 5.5, we know that for all $\widetilde{P_1}, \widetilde{P_2} \geq 0$, $D_L(\widetilde{P_1}, \widetilde{P_2}) \geq 1$. Thus, by (10.136) we can conclude

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq 1 \geq \frac{9}{10}(\frac{9}{10\sqrt{2}})\Sigma_1$$

Finally, by (i),(ii),(iii),

$$\begin{aligned} D &\geq \min(\frac{1}{1+\sqrt{2}}, \frac{1}{\frac{2}{\sqrt{2}}+2}\frac{9}{10}, \frac{9}{10}(\frac{9}{10\sqrt{2}}))\Sigma_1 \\ &= \frac{1}{\frac{2}{\sqrt{2}}+2}\frac{9}{10}\Sigma_1 \\ &\geq 0.26360...\Sigma_1. \end{aligned}$$

Since by Lemma 5.5 we already know $D_L(\widetilde{P_1}, \widetilde{P_2}) \geq 1$, the statement (e) of the corollary is true. □

Proof of Proposition 5.3 of Page 253. Like the proof of Proposition 5.1, we define the subscript max as $argmax_{i\in\{1,2\}}\widetilde{P}_i$, and write $D_{\sigma_{v1}}(\cdot), D_{\sigma_{v2}}(\cdot), D_{\sigma_{vmax}}(\cdot)$ as $D_{v1}(\cdot), D_{v2}(\cdot), D_{vmax}(\cdot)$ respectively.

By the same argument as the proof of Proposition 5.1, it is enough to show that there exists $c \leq 10^6$ such that for all $\widetilde{P_1}, \widetilde{P_2} \geq 0$, $\min(D_{\sigma 1}(c\widetilde{P_1}), D_{\sigma 2}(c\widetilde{P_2})) \leq c \cdot D_L(\widetilde{P_1}, \widetilde{P_2})$.

In the proof, we first divide the cases based on Σ_1, Σ_2, and then based on $\widetilde{P_1}, \widetilde{P_2}$. Here, we know $\Sigma_1 \leq \Sigma_2$ since $\sigma_1 \leq \sigma_2$. Using this, we can reduce the cases.

(i) When $\Sigma_1 \leq 40$, $\Sigma_2 \leq 40$

(i-i) When $\max(\widetilde{P_1}, \widetilde{P_2}) \geq \frac{1}{40}$

Lower bound: By Corollary 5.6 (e), we have

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq 1$$

If $\max(\widetilde{P_1}, \widetilde{P_2}) \geq \frac{1}{40}$ and $\Sigma_{max} \geq \frac{1}{1-a^2}$

Upper bound: By Corollary 5.3 (5.10), we have

$$(D_{\sigma max}(P_{max}), P_{max}) \leq (\frac{1}{1-a^2}, 0) \leq (\Sigma_{max}, 0) \leq (40, 0).$$

If $\max(\widetilde{P_1}, \widetilde{P_2}) \geq \frac{1}{40}$ and $\Sigma_{max} \leq \frac{1}{1-a^2}$

Upper bound: By plugging $t = \frac{1}{\max(1,\Sigma_{max})}$ into Corollary 5.3 (5.11), we have

$$\begin{aligned}(D_{\sigma max}(P_{max}), P_{max}) &\leq (2\max(1, \Sigma_{max}), \frac{1}{\max(1, \Sigma_{max})}) \\ &\leq (2 \cdot 40, 1)(\because \text{In (i), we assumed } \Sigma_1 \leq 40, \Sigma_2 \leq 40)\end{aligned}$$

Ratio: c is upper bounded by

$$c \leq 2 \cdot 40.$$

(i-ii) When $\frac{1-a^2}{20} \leq \max(\widetilde{P_1}, \widetilde{P_2}) \leq \frac{1}{40}$

Lower bound: By Corollary 5.6 (c), we have

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq \frac{0.001201}{\max(\widetilde{P_1}, \widetilde{P_2})} + 1. \tag{10.137}$$

If $\Sigma_{max} \geq \frac{1}{1-a^2}$

Upper bound: By Corollary 5.3 (5.10), we have

$$(D_{\sigma max}(P_{max}), P_{max}) \leq (\frac{1}{1-a^2}, 0) \leq (\Sigma_{max}, 0) \leq (40, 0)$$

If $\Sigma_{max} \leq \frac{1}{1-a^2}$ and $1 - a^2 \leq \max(\widetilde{P_1}, \widetilde{P_2}) \leq \frac{1}{40}$

Since we assume $\Sigma_1 \leq 40$, $\Sigma_2 \leq 40$ in (i), we have $1 - a^2 \leq \widetilde{P_{max}} \leq \frac{1}{40} \leq \frac{1}{\max(1,\Sigma_{max})}$. Thus, we can plug $t = \widetilde{P_{max}}$ to Corollary 5.3 (5.11), and conclude

$$(D_{\sigma max}(P_{max}), P_{max}) \leq (\frac{2}{\widetilde{P_{max}}}, \widetilde{P_{max}}).$$

If $\Sigma_{max} \leq \frac{1}{1-a^2}$ and $\frac{1-a^2}{20} \leq \max(\widetilde{P_1}, \widetilde{P_2}) \leq 1 - a^2$

In this case, the lower bound of (10.137) can be further lower bounded by

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq \frac{0.001201}{1-a^2} + 1.$$

Upper bound: By Corollary 5.3 (5.10), we have

$$(D_{\sigma max}(P_{max}), P_{max}) \leq (\frac{1}{1-a^2}, 0).$$

Ratio: c is upper bounded by

$$c \leq \frac{2}{0.001201} < 2000.$$

(i-iii) When $\max(\widetilde{P_1}, \widetilde{P_2}) \le \frac{1-a^2}{20}$

Lower bound: By Corollary 5.6 (d), we have

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \ge \frac{0.0869}{1-a^2} + 1$$

Upper bound: By Corollary 5.3 (5.10), we have

$$(D_{\sigma max}(P_{max}), P_{max}) \le (\frac{1}{1-a^2}, 0).$$

Ratio: c is upper bounded by

$$c \le \frac{1}{0.0869} < 12.$$

(ii) When $\Sigma_1 \le 40 \le \Sigma_2$

(ii-i) When $\widetilde{P_1} \ge \frac{1}{40}$

Lower bound: By Corollary 5.6 (e), we have

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \ge 1.$$

If $\widetilde{P_1} \ge \frac{1}{40}$ and $\Sigma_1 \ge \frac{1}{1-a^2}$

Upper bound: By Corollary 5.3 (5.10), we have

$$(D_{\sigma 1}(P_1), P_1) \le (\frac{1}{1-a^2}, 0) \le (\Sigma_1, 0) \le (40, 0).$$

If $\widetilde{P_1} \ge \frac{1}{40}$ and $\Sigma_1 \le \frac{1}{1-a^2}$

Upper bound: By plugging $t = \frac{1}{\max(1,\Sigma_1)}$ into the equation (5.11) of Corollary 5.3, we have

$$\begin{aligned}(D_{\sigma 1}(P_1), P_1) &\le (2\max(1, \Sigma_1), \frac{1}{\max(1, \Sigma_1)}) \\ &\le (2 \cdot 40, 1) (\because \text{In (ii), we assumed } \Sigma_1 \le 40)\end{aligned}$$

Ratio: c is upper bounded by

$$c \le 2 \cdot 40.$$

(ii-ii) When $\frac{1}{\Sigma_2} \le \widetilde{P_1} \le \frac{1}{40}$

Lower bound: By Corollary 5.6 (b), we have

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \ge \frac{0.009131}{\widetilde{P_1}} + 1. \tag{10.138}$$

If $\Sigma_1 \ge \frac{1}{1-a^2}$

Upper bound: By Corollary 5.3 (5.10), we have

$$(D_{\sigma 1}(P_1), P_1) \le (\frac{1}{1-a^2}, 0) \le (\Sigma_1, 0) \le (40, 0)$$

If $\Sigma_1 \leq \frac{1}{1-a^2}$ and $1 - a^2 \leq \widetilde{P_1} \leq \frac{1}{40}$

Upper bound: Since $1 - a^2 \leq \widetilde{P_1} \leq \frac{1}{40} \leq \frac{1}{\max(1,\Sigma_1)}$, we can plug $t = \widetilde{P_1}$ into Corollary 5.3 (5.11). Thus, we have

$$(D_{\sigma 1}(P_1), P_1) \leq (\frac{2}{\widetilde{P_1}}, \widetilde{P_1}).$$

If $\Sigma_1 \leq \frac{1}{1-a^2}$ and $\frac{1}{\Sigma_2} \leq \widetilde{P_1} \leq 1 - a^2$

In this case, the lower bound of (10.138) can be further lower bounded by

$$D_L(P_1, P_2) \geq \frac{0.009131}{1-a^2} + 1.$$

Upper bound: By Corollary 5.3 (5.10), we have

$$(D_{\sigma 1}(P_1), P_1) \leq (\frac{1}{1-a^2}, 0).$$

Ratio: c is upper bounded by

$$c \leq \frac{2}{0.009131} < 220.$$

(ii-iii) When $\widetilde{P_1} \leq \frac{1}{\Sigma_2}$ and $\max(\widetilde{P_1}, \widetilde{P_2}) = \widetilde{P_2} > \frac{1}{\Sigma_2}$

Lower bound: By Corollary 5.6 (a), we have

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq 0.009131\Sigma_2 + 1. \tag{10.139}$$

If $\Sigma_2 \geq \frac{1}{1-a^2}$, $\widetilde{P_1} \leq \frac{1}{\Sigma_2}$ and $\max(\widetilde{P_1}, \widetilde{P_2}) > \frac{1}{\Sigma_2}$

The lower bound of (10.139) can be further lower bonded by

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq 0.009131\Sigma_2 + 1 \geq 0.009131\frac{1}{1-a^2} + 1.$$

Upper bound: By Corollary 5.3 (5.10), we have

$$(D_{\sigma 2}(P_2), P_2) \leq (\frac{1}{1-a^2}, 0).$$

If $\Sigma_2 \leq \frac{1}{1-a^2}$, $\widetilde{P_1} \leq \frac{1}{\Sigma_2}$ and $\max(\widetilde{P_1}, \widetilde{P_2}) > \frac{1}{\Sigma_2}$

Upper bound: Since we assumed $\Sigma_2 \geq 40$ in (ii), $\max(1, \Sigma_2) = \Sigma_2$. Thus, we can plug $t = \frac{1}{\Sigma_2}$ into (5.11) of Corollary 5.3, and conclude

$$(D_{\sigma 2}(P_2), P_2) \leq (2\Sigma_2, \frac{1}{\Sigma_2}).$$

Ratio: c is upper bounded by

$$c \leq \frac{2}{0.009131} < 220.$$

(ii-iv) When $\widetilde{P_1} \leq \frac{1}{\Sigma_2}$ and $\frac{1-a^2}{20} \leq \max(\widetilde{P_1}, \widetilde{P_2}) \leq \frac{1}{\Sigma_2}$

Lower bound: Since $\frac{1-a^2}{20} \leq \max(\widetilde{P_1}, \widetilde{P_2}) \leq \frac{1}{\Sigma_2} \leq \frac{1}{40}$, by Corollary 5.6 (c) we have

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq \frac{0.001201}{\max(\widetilde{P_1}, \widetilde{P_2})} + 1. \tag{10.140}$$

If $\Sigma_{max} \geq \frac{1}{1-a^2}$, $\widetilde{P_1} \leq \frac{1}{\Sigma_2}$ and $\frac{1-a^2}{20} \leq \max(\widetilde{P_1}, \widetilde{P_2}) \leq \frac{1}{\Sigma_2}$
In this case, the lower bound of (10.140) can be further lower bounded by

$$\begin{aligned} D_L(\widetilde{P_1}, \widetilde{P_2}) &\geq \frac{0.001201}{\max(\widetilde{P_1}, \widetilde{P_2})} + 1 \geq 0.001201\Sigma_2 + 1 \\ &\geq 0.001201\Sigma_{max} + 1 \geq \frac{0.001201}{1-a^2} + 1. \end{aligned}$$

Upper bound: By Corollary 5.3 (5.10), we have

$$(D_{\sigma max}(P_{max}), P_{max}) \leq (\frac{1}{1-a^2}, 0).$$

If $\Sigma_{max} \leq \frac{1}{1-a^2}$, $\widetilde{P_1} \leq \frac{1}{\Sigma_2}$ and $\frac{1}{\max(1,\Sigma_{max})} < \max(\widetilde{P_1}, \widetilde{P_2}) \leq \frac{1}{\Sigma_2}$
This case never happens since $\Sigma_2 \geq \max(1, \Sigma_{max})$.
If $\Sigma_{max} \leq \frac{1}{1-a^2}$, $\widetilde{P_1} \leq \frac{1}{\Sigma_2}$ and $1 - a^2 \leq \max(\widetilde{P_1}, \widetilde{P_2}) \leq \frac{1}{\max(1,\Sigma_{max})}$
Upper bound: By plugging $t = \widetilde{P_{max}}$ into (5.11) of Corollary 5.3, we have

$$(D_{\sigma max}(P_{max}), P_{max}) \leq (\frac{2}{\widetilde{P_{max}}}, \widetilde{P_{max}}).$$

If $\Sigma_{max} \leq \frac{1}{1-a^2}$, $\widetilde{P_1} \leq \frac{1}{\Sigma_2}$ and $\frac{1-a^2}{20} \leq \max(\widetilde{P_1}, \widetilde{P_2}) \leq 1 - a^2$
In this case, the lower bound of (10.140) can be further lower bounded by

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq \frac{0.0012011}{1-a^2} + 1.$$

Upper bound: By Corollary 5.3 (5.10), we have

$$(D_{\sigma max}(P_{max}), P_{max}) \leq (\frac{1}{1-a^2}, 0).$$

Ratio: c is upper bounded by

$$c \leq \frac{2}{0.0012011} \leq 1700.$$

(ii-v) When $\widetilde{P_1} \leq \frac{1}{\Sigma_2}$ and $\max(\widetilde{P_1}, \widetilde{P_2}) \leq \frac{1-a^2}{20}$
Lower bound: By Corollary 5.6 (d), we have

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq \frac{0.0869}{1-a^2} + 1.$$

Upper bound: By Corollary 5.3 (5.10), we have

$$(D_{\sigma max}(P_{max}), P_{max}) \leq (\frac{1}{1-a^2}, 0).$$

Ratio: c is upper bounded by

$$c \leq \frac{1}{0.0869} \leq 12.$$

(iii) When $40 \leq \Sigma_1 \leq \Sigma_2$

(iii-i) When $\widetilde{P_1} \geq \frac{1}{\Sigma_1}$

Lower bound: By Corollary 5.6 (e), we have

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq 0.2636\Sigma_1.$$

If $\widetilde{P_1} \geq \frac{1}{\Sigma_1}$ and $\Sigma_1 \geq \frac{1}{1-a^2}$

Upper bound: By Corollary 5.3 (5.10), we have

$$(D_{\sigma 1}(\widetilde{P_1}), \widetilde{P_1}) \leq (\frac{1}{1-a^2}, 0) \leq (\Sigma_1, 0).$$

If $\widetilde{P_1} \geq \frac{1}{\Sigma_1}$ and $\Sigma_1 \leq \frac{1}{1-a^2}$

Upper bound: By plugging $t = \frac{1}{\Sigma_1}$ into (5.11) of Corollary 5.3, we have

$$(D_{\sigma 1}(P_1), P_1) \leq (2\Sigma_1, \frac{1}{\Sigma_1}).$$

Ratio: c is upper bounded by

$$c \leq \frac{2}{0.2636} < 8.$$

(iii-ii) When $\frac{1}{\Sigma_2} \leq \widetilde{P_1} \leq \frac{1}{\Sigma_1}$

Lower bound: Since $\frac{1}{\Sigma_2} \leq \widetilde{P_1} \leq \frac{1}{\Sigma_1} \leq \frac{1}{40}$, by Corollary 5.6 (b)

$$we have D_L(\widetilde{P_1}, \widetilde{P_2}) \geq \frac{0.009131}{\widetilde{P_1}} + 1. \tag{10.141}$$

If $\Sigma_1 \geq \frac{1}{1-a^2}$ and $\frac{1}{\Sigma_2} \leq \widetilde{P_1} \leq \frac{1}{\Sigma_1}$

In this case, the lower bound of (10.141) can be further lower bounded by

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq 0.009131\Sigma_1 + 1.$$

Upper bound: By Corollary 5.3 (5.10), we have

$$(D_{\sigma 1}(P_1), P_1) \leq (\frac{1}{1-a^2}, 0) \leq (\Sigma_1, 0).$$

If $\Sigma_1 \leq \frac{1}{1-a^2}$ and $1 - a^2 \leq \widetilde{P_1} \leq \frac{1}{\Sigma_1}$

Upper bound: By plugging $t = \widetilde{P_1}$ into (5.11) of Corollary 5.3, we have

$$(D_{\sigma 1}(P_1), P_1) \leq (\frac{2}{\widetilde{P_1}}, \widetilde{P_1}).$$

If $\Sigma_1 \leq \frac{1}{1-a^2}$ and $\frac{1}{\Sigma_2} \leq \widetilde{P_1} \leq 1 - a^2$

In this case, the lower bound (10.141) can be further lower bounded by

$$D_L(\widetilde{P_1}, \widetilde{P_2}) \geq \frac{0.009131}{1-a^2} + 1.$$

Upper bound: By Corollary 5.3 (5.10), we have

$$(D_{\sigma 1}(P_1), P_1) \leq (\frac{1}{1-a^2}, 0).$$

Ratio: c is upper bounded by

$$c \leq \frac{2}{0.009131} < 220.$$

(iii-iii) When $\widetilde{P_1} \leq \frac{1}{\Sigma_2}$ and $\max(\widetilde{P_1}, \widetilde{P_2}) > \frac{1}{\Sigma_2}$

Compared to the case (ii-iii), the only difference is the condition on Σ_1. Moreover, since the argument of (ii-iii) does not depend on the condition on Σ_1, we can get the same bound on c following the same argument as (ii-iii).

(iii-iv) When $\widetilde{P_1} \leq \frac{1}{\Sigma_2}$ and $\frac{1-a^2}{20} \leq \max(\widetilde{P_1}, \widetilde{P_2}) \leq \frac{1}{\Sigma_2}$

Compared to the case (ii-iv), the only difference is the condition on Σ_1. Moreover, since the argument of (ii-iv) does not depend on the condition on Σ_1, we can get the same bound on c following the same argument as (ii-iv).

(iii-v) When $\widetilde{P_1} \leq \frac{1}{\Sigma_2}$ and $\max(\widetilde{P_1}, \widetilde{P_2}) \leq \frac{1-a^2}{20}$

Compared to the case (ii-v), the only difference is the condition on Σ_1. Moreover, since the argument of (ii-v) does not depend on the condition on Σ_1, we can get the same bound on c following the same argument as (ii-v). □

Bibliography

[1] R. Ahlswede, N. Cai, S.-Y.R. Li, and R.W. Yeung. Network information flow. *IEEE Transactions on Information Theory*, 46(4):1204–1216, July 2000.

[2] A. Amaudruz and C. Fragouli. Combinatorial algorithms for wireless information flow. *Proceedings of ACM-SIAM Symposium on Discrete Algorithms (SODA09)*, Jan 2009.

[3] B. Anderson. Transfer function matrix description of decentralized fixed modes. *IEEE Trans. Automat. Contr.*, 6, December 1982.

[4] B. Anderson and D. Clements. Algebraic characterization of fixed modes in decentralized control. *Automatica*, 17(5):703–712, September 1981.

[5] B. Anderson and J. Moore. Time-varying feedback laws for decentralized control. *IEEE Trans. Automat. Contr.*, 26(5):1133–1139, October 1981.

[6] S. Avestimehr, S. Diggavi, and D. Tse. Wireless network information flow: A deterministic approach. *IEEE Trans. Inform. Theory*, 57(4):1872 –1905, April 2011.

[7] Baglietto, M. Parisini, and R. T. Zoppoli. Numerical solutions to the Witsenhausen counterexample by approximating networks. *IEEE Trans. Automat. Contr.*, 46(9), 2001.

[8] J Baillieul. Feedback designs for controlling device arrays with communication channel bandwidth constraints. In *ARO Workshop on Smart Structures, Pennsylvania State Univ*, pages 16–18, 1999.

[9] J Baillieul. Feedback coding for information-based control: operating near the data-rate limit. In *Decision and Control, 2002, Proceedings of the 41st IEEE Conference on*, volume 3, pages 3229–3236. IEEE, 2002.

[10] R. Bansal and T. Basar. Stochastic teams with nonclassical information revisited: When is an affine law optimal? *Automatic Control, IEEE Transactions on*, 32(6):554 – 559, jun 1987.

[11] D. Bertsekas. *Dynamic Programming and Optimal Control.* Athena Scientific, third edition, 2005.

[12] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization.* Cambridge University Press, 2004.

[13] Miran Brezocnik, Joze Balic, and Zmago Brezocnik. Emergence of intelligence in next-generation manufacturing systems. *Robotics and computer-integrated manufacturing*, 19(1):55–63, 2003.

[14] V. Cadambe and S. Jafar. Interference alignment and degrees of freedom of the k-user interference channel. *Information Theory, IEEE Transactions on*, 54(8):3425–3441, 2008.

[15] E.J. Candes, J.Romberg, and T.Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2), February 2006.

[16] Alvaro A Cardenas, Saurabh Amin, and Shankar Sastry. Secure control: Towards survivable cyber-physical systems. In *Distributed Computing Systems Workshops, 2008. ICDCS'08. 28th International Conference on*, pages 495–500. IEEE, 2008.

[17] Chi-Tsong Chen. *Linear System Theory and Design.* Oxford University Press, New York, NY, third edition, 1999.

[18] Chiranjib Choudhuri and Urbashi Mitra. On witsenhausen's counterexample: The asymptotic vector case. In *Information Theory Workshop (ITW), 2012 IEEE*, pages 162 –166, sept. 2012.

[19] J. P. Corfmat and A. S. Morse. Decentralized control of linear multivariable systems. *Automatica*, 12, September 1976.

[20] Max Costa. Writing on dirty paper (corresp.). *Information Theory, IEEE Transactions on*, 29(3):439–441, 1983.

[21] T. M. Cover and J. A. Thomas. *Elements of Information Theory.* Wiley, 2006.

[22] T.M. Cover and S. Pombra. Gaussian feedback capacity. *Information Theory, IEEE Transactions on*, 35(1):37 –43, jan 1989.

[23] E. Davison and T. Chang. Decentalized stabilization and pole assignment for general proper systems. *IEEE Trans. Automat. Contr.*, 35(6):652–664, June 1990.

[24] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications.* Springer, second edition, 1998.

[25] D.L. Donoho. Compressed sensing. *Manuscript.*

[26] N. Elia. When Bode meets Shannon: Control-oriented feedback communication schemes. *IEEE Trans. Automat. Contr.*, 49(9):1477–1488, September 2004.

[27] N. Elia. Remote stabilization over fading channels. *Systems and Control Letters*, 54(3):239–249, March 2005.

[28] P. Elias, A. Feinstein, and C. Shannon. A note on the maximum flow through a network. *IRE Transactions on Information Theory*, 2(4):117–119, December 1956.

[29] R.H. Etkin, D.N.C. Tse, and Hua Wang. Gaussian interference channel capacity to within one bit. *Information Theory, IEEE Transactions on*, 54(12):5534 –5562, dec. 2008.

[30] W. Feller. *An introduction to probability theory and its applications.* Wiley, New York, NY, sixth edition, 1957.

[31] L. R. Ford and D. R. Fulkerson. Maximal flow through a network. *Canadian Journal of Mathematics*, 8:399–404, 1956.

[32] C. W. Gardiner. *Handbook of Stochastic Methods.* Springer, third edition, November 2003.

[33] E. Garone, B. Sinopoli, A. Goldsmith, and A. Casavola. Lqg control for distributed systems over tcp-like erasure channels. *Proceedings of IEEE Conference on Decision and Control*, 2007.

[34] M. Gastpar, B. Rimoldi, and M. Vetterli. To code, or not to code: lossy source-channel communication revisited. *Information Theory, IEEE Transactions on*, 49(5):1147 – 1158, may 2003.

[35] B. Gelbaum and J. Olmsted. *Counterexamples in Analysis.* Dover, 1964.

[36] M. Goemans, S. Iwata, and R. Zenklusen. An algorithmic framework for wireless information flow. *Proceedings of the 47th Annual Allerton Conference on Communication, Control and Computing, Monticello, IL, USA*, 2009.

[37] P. Grover, S. Y. Park, and A. Sahai. The finite-dimensional Witsenhausen counterexample. *IEEE Trans. Automat. Contr.*, To appear.

[38] P. Grover, A.B. Wagner, and A. Sahai. Information embedding meets distributed control. In *Information Theory Workshop (ITW), 2010 IEEE*, pages 1 –5, jan. 2010.

[39] Pulkit Grover and Anant Sahai. Witsenhausen's counterexample as assisted interference suppression. *International Journal of Systems, Control and Communications*, 2(1):197–237, 2010.

[40] Piyush Gupta and Panganmala R Kumar. The capacity of wireless networks. *Information Theory, IEEE Transactions on*, 46(2):388–404, 2000.

[41] Rajesh Gupta. Programming models and methods for spatiotemporal actions and reasoning in cyber-physical systems. In *NSF Workshop on CPS*, 2006.

[42] Vijay Gupta, Babak Hassibi, and Richard M Murray. Optimal lqg control across packet-dropping links. *Systems & Control Letters*, 56(6):439–446, 2007.

[43] Godfrey H Hardy, John E Littlewood, and George Pólya. *Inequalities, Cambridge Mathematical Library*. Cambridge University Press, Cambridge, 1988.

[44] Joao P Hespanha, Payam Naghshtabrizi, and Yonggang Xu. A survey of recent results in networked control systems. *Proceedings of the IEEE*, 95(1):138–162, 2007.

[45] Yu-Chi Ho, M. P. Kastner, and E. Wong. Teams, signaling, and information theory. *IEEE Trans. Automat. Contr.*, 23(2), April 1978.

[46] D. Hochbaum. IEOR 269 integer programming and combinatorial optimization lecture note. *http://www.ieor.berkeley.edu/∼hochbaum/*.

[47] http://www.its.dot.gov/newsletter/january2013.htm.

[48] J. Karlsson, A. Gattami, T.J. Oechtering, and M. Skoglund. Iterative source-channel coding approach to Witsenhausen's counterexample. In *American Control Conference (ACC), 2011*, pages 5348 –5353, 29 2011-july 1 2011.

[49] M. Kim and M. Medard. Algebraic network coding approach to deterministic wireless relay networks. *arXiv:1001.4431v2*, Jun 2010.

[50] Young-Han Kim, A. Sutivong, and T.M. Cover. State amplification. *Information Theory, IEEE Transactions on*, 54(5):1850 –1859, may 2008.

[51] Jacob Klimstra and Markus Hotakainen. *Smart power generation*. Avain, 2011.

[52] R. Koetter and M. Medard. An algebraic approach to network coding. *IEEE ACM Transations On Networking*, 11:782–795, 2003.

[53] Steven G Krantz and Harold R Parks. *A primer of real analytic functions*. Birkhäuser Boston, 2002.

[54] K. Kuipers and H. Niederreitter. *Uniform Distribution of Sequences*. Wiley, 1974.

[55] P. R. Kumar and P. Varaiya. *Stochastic Systems: Estimation, Identification, and Adaptive Control*. Prentice Hall, New Jersey, 1986.

[56] Javad Lavaei and Somayeh Sojoudi. Time complexity of decentralized fixed-mode verification. *Automatic Control, IEEE Transactions on*, 55(4):971–976, 2010.

[57] Edward A Lee. Cyber physical systems: Design challenges. In *Object Oriented Real-Time Distributed Computing (ISORC), 2008 11th IEEE International Symposium on*, pages 363–369. IEEE, 2008.

[58] Insup Lee and Oleg Sokolsky. Medical cyber physical systems. In *Design Automation Conference (DAC), 2010 47th ACM/IEEE*, pages 743–748. IEEE, 2010.

[59] J.T. Lee, E. Lau, and Yu-Chi Ho. The Witsenhausen counterexample: a hierarchical search approach for nonconvex optimization problems. *Automatic Control, IEEE Transactions on*, 46(3):382 –397, mar 2001.

[60] Laurent Lessard and Sanjay Lall. A state-space solution to the two-player decentralized optimal control problem. In *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*, pages 1559 –1564, sept. 2011.

[61] Na Li, J.R. Marden, and J.S. Shamma. Learning approaches to the Witsenhausen counterexample from a view of potential games. In *Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on*, pages 157 –162, dec. 2009.

[62] S.-Y. R. Li and R. W. Yeung. Network information flow multiple sources. *Proc. 2001 IEEE Int. Symp. Information Theory.*

[63] Mohammad Ali Maddah-Ali, Abolfazl Seyed Motahari, and Amir K Khandani. Communication over mimo x channels: Interference alignment, decomposition, and performance analysis. *Information Theory, IEEE Transactions on*, 54(8):3457–3470, 2008.

[64] N.C. Martins and M.A. Dahleh. Feedback control in the presence of noisy channels: Bode-like fundamental limitations of performance. *Automatic Control, IEEE Transactions on*, 53(7):1604 –1615, aug. 2008.

[65] M. Medard. Personal communication.

[66] P. Minero, M. Franceschetti, S. Dey, and G.N. Nair. Data rate theorem for stabilization over time-varying feedback channels. *Automatic Control, IEEE Transactions on*, 54(2):243 –255, feb. 2009.

[67] Sanjoy K Mitter and Nigel J Newton. Information and entropy flow in the kalman–bucy filter. *Journal of Statistical Physics*, 118(1-2):145–176, 2005.

[68] Sanjoy K. Mitter and Anant Sahai. Information and control: Witsenhausen revisited. *Learning, Control and Hybrid Systems: Lecture Notes in Control and Information Sciences 241*, pages 281–293, 1999.

[69] G. N. Nair and R. J. Evans. Communication-limited stabilization of linear systems. In *Proceedings of the 39th IEEE Conference on Decision and control*, pages 1005–1010, Sydney, Australia, December 2000.

[70] Girish N Nair and Robin J Evans. Exponential stabilisability of finite-dimensional linear systems with limited data rates. *Automatica*, 39(4):585–593, 2003.

[71] Girish N Nair, Fabio Fagnani, Sandro Zampieri, and Robin J Evans. Feedback control under data rate constraints: An overview. *Proceedings of the IEEE*, 95(1):108–137, 2007.

[72] Ashutosh Nayyar, Aditya Mahajan, and Demosthenis Teneketzis. Decentralized stochastic control with partial sharing information structures: a common information approach. *IEEE Transactions on Automatic Control, to appear*, 2013.

[73] B. Nazer and M. Gastpar. Computation over multiple-access channels. *IEEE Trans. Inform. Theory*, 53(10):3498–3516, October 2007.

[74] Bobak Nazer and Michael Gastpar. Computation over multiple-access channels. *Information Theory, IEEE Transactions on*, 53(10):3498–3516, 2007.

[75] A. V. Oppenheim, A. S. Willsky, and S. H. Nawab. *Signal and Systems.* Springer-Verlag, New York, NY, second edition, 1991.

[76] M. Pajic, S. Sundaram, G.J. Pappas, and R. Mangharam. Topological conditions for wireless control networks. In *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*, pages 2353 –2360, dec. 2011.

[77] Christos H Papadimitriou and John Tsitsiklis. Intractable problems in control theory. *SIAM Journal on Control and Optimization*, 24(4):639–654, 1986.

[78] S. Y. Park, P. Grover, and A. Sahai. A constant-factor approximately optimal solution to the witsenhausen counterexample. In *Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on*, pages 2881 –2886, dec. 2009.

[79] P.P.Vaidyanathan and V.C.Liu. Efficient reconstruction of band-limited sequences from nonuniformly decimated versions by use of polyphase filter banks. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 38(11), November 1990.

[80] R. Radner. Team decision problems. *The Annals of Mathematical Statistics*, 33(3):pp. 857–881.

[81] G. Ranade and A. Sahai. Implicit communication in multiple-access settings. In *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*, pages 998 –1002, 31 2011-aug. 5 2011.

[82] R.J.Evans and I.M.Isaacs. Generalized vandermonde determinants and roots of unity of prime order. *Proceedings of the American Mathematical Society*, 58, July 1976.

[83] M. Rotkowitz. On information structures, convexity, and linear optimality. In *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, pages 1642 –1647, dec. 2008.

[84] M. Rotkowitz and S. Lall. A characterization of convex problems in decentralized control. *Automatic Control, IEEE Transactions on*, 50(12):1984 –1996, dec. 2005.

[85] A. Sahai. Why do block length and delay behave differently if feedback is present? *Information Theory, IEEE Transactions on*, 54(5):1860 –1886, may 2008.

[86] A. Sahai and S. Mitter. The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link - Part I: Scalar system. *IEEE Trans. Inform. Theory*, 52(8):3369–3395, August 2006.

[87] Anant Sahai. *Anytime Information Theory.* PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, 2001.

[88] Jr. Sandell, N., P. Varaiya, M. Athans, and M. Safonov. Survey of decentralized control methods for large scale systems. *Automatic Control, IEEE Transactions on*, 23(2):108 – 128, apr 1978.

[89] J. P. M. Schalkwijk and T. Kailath. A coding scheme for additive noise channels with feedback – I: No bandwidth constraint. *IEEE Trans. Inform. Theory*, 12(2):172–182, April 1966.

[90] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. Sastry. Foundations of control and estimation over lossy networks. *Proceedings of the IEEE*, 95(1):163–187, January 2007.

[91] A. Schrijver. A combinatorial algorithm minimizing submodular functions in strongly polynomial time. *Journal of Combinatorial Theory, Series B*, 80:346–355, 2000.

[92] P. Shah and P.A. Parrilo. A partial order approach to decentralized control. In *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, pages 4351 –4356, dec. 2008.

[93] C. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423,623–656, October 1948.

[94] Claude E Shannon. Communication theory of secrecy systems. *Bell system technical journal*, 28(4):656–715, 1949.

[95] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. Jordan, and S. Sastry. Kalman filtering with intermittent observations. *IEEE Trans. Automat. Contr.*, 49(9):1453–1464, September 2004.

[96] Paulo Tabuada. Cyber-physical systems: Position paper. In *NSF Workshop on Cyber-Physical Systems*, 2006.

[97] S. Tatikonda and S. K. Mitter. Control under communication constraints. *IEEE Trans. Automat. Contr.*, 49(7):1056–1068, July 2004.

[98] Simon Thorpe, Denis Fize, Catherine Marlot, et al. Speed of processing in the human visual system. *nature*, 381(6582):520–522, 1996.

[99] D. Tse and P. Viswanath. *Fundamentals of Wireless Communication.* Cambridge University Press, 2005.

[100] John N Tsitsiklis and Vincent D Blondel. The lyapunov exponent and joint spectral radius of pairs of matrices are hardwhen not impossibleto compute and to approximate. *Mathematics of Control, Signals, and Systems (MCSS)*, 10(1):31–40, 1997.

[101] R. Varga. *Matrix Iterative Analysis.* Springer, second edition, 2002.

[102] Chih-Chun Wang and Ness B Shroff. Pairwise intersession network coding on directed networks. *Information Theory, IEEE Transactions on*, 56(8):3879–3900, 2010.

[103] I-Hsiang Wang, Sudeep U Kamath, and David NC Tse. Two unicast information flows over linear deterministic networks. In *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*, pages 2462–2466. IEEE, 2011.

[104] Shih-Ho Wang and E Davison. On the stabilization of decentralized control systems. *Automatic Control, IEEE Transactions on*, 18(5):473–478, 1973.

[105] N. Wiener. *Cybernetics, or Communication and Control in the Animal and the Machine.* Cambridge MIT Press, 1948.

[106] Norbert Wiener. *Cybernetics: or the Control and Communication in the Animal and the Machine*, volume 25. The MIT press, 1961.

[107] Norbert Wiener. *Extrapolation, interpolation, and smoothing of stationary time series: with engineering applications*, volume 8. MIT press, 1964.

[108] H. S. Witsenhausen. A counterexample in stochastic optimum control. *SIAM Journal on Control*, 6(1):131–147, January 1968.

[109] H. S. Witsenhausen. Separation of estimation and control for discrete time systems. *Proceedings of the IEEE*, 59(11):1557–1566, November 1971.

[110] Yihong Wu and Sergio Verdu. Witsenhausen's counterexample: A view from optimal transport theory. In *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*, pages 5732 –5737, dec. 2011.

[111] Feng Xia and Jianhua Ma. Building smart communities with cyber-physical systems. In *Proceedings of 1st international symposium on From digital footprints to social and community intelligence*, pages 1–6. ACM, 2011.

[112] S. M. S. Tabatabaei Yazdi and S. A. Savari. A max-flow/min-cut algorithm for a class of wireless networks. *the Proceedings of the ACM-SIAM Symposium on Discrete Algorithms (SODA10)*, Jan 2010.

[113] M. Yilin and B. Sinopoli. A characterization of the critical value for Kalman filtering with intermittent observations. *Proceedings of the 47th IEEE Conference on Decision and Control CDC*, 2008.

[114] D. Youla, H. Jabr, and Jr. Bongiorno, J. Modern Wiener-Hopf design of optimal controllers–part ii: The multivariable case. *Automatic Control, IEEE Transactions on*, 21(3):319 – 338, jun 1976.

[115] Wei Yu and Raymond Lui. Dual methods for nonconvex spectrum optimization of multicarrier systems. *Communications, IEEE Transactions on*, 54(7):1310–1322, 2006.

[116] S. Yuksel. Decentralized computation and communication in stabilization of distributed control systems. *Information Theory and Applications Workshop*, 2009.

[117] S. Yuksel. Stochastic nestedness and the belief sharing information pattern. *Automatic Control, IEEE Transactions on*, 54(12):2773 –2786, dec. 2009.

[118] S Yuksel and T Basar. Minimum rate coding for lti systems over noiseless channels. *Automatic Control, IEEE Transactions on*, 51(12):1878–1887, 2006.

[119] S. Yuksel and T. Basar. Optimal signaling policies for decentralized multicontroller stabilizability over communication channels. *Automatic Control, IEEE Transactions on*, 52(10):1969 –1974, oct. 2007.

[120] S Yuksel and T Basar. Control over noisy forward and reverse channels. *Automatic Control, IEEE Transactions on*, 56(5):1014–1029, 2011.