Active systems with uncertain parameters: an information-theoretic perspective

Gireeja Ranade

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2015-43
http://www.eecs.berkeley.edu/Pubs/TechRpts/2015/EECS-2015-43.html

May 1, 2015

Active Systems with Uncertain Parameters: an Information-Theoretic Perspective

by

Gireeja Vishnu Ranade

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Electrical Engineering and Computer Sciences

and the Designated Emphasis

in

Communication, Computation and Statistics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Anant Sahai, Chair
Professor David Tse
Professor Pravin Varaiya
Professor Adityanand Guntuboyina

Fall 2014

Active Systems with Uncertain Parameters: an Information-Theoretic Perspective

Abstract

Active Systems with Uncertain Parameters: an Information-Theoretic Perspective

by

Gireeja Vishnu Ranade

Doctor of Philosophy in Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Anant Sahai, Chair

The sister fields of control and communication share many common goals. While communication aims to reduce uncertainty by communicating information about the state of the world, control aims to reduce uncertainty by moving the state of the world to a known point. Furthermore, transmitters must communicate over the unreliability in the communication channel, while controllers must overcome unreliability in the sensing and actuation channels of the system to stabilize it. Extensive work in information theory has provided a framework with which we can understand the fundamental limits on the communication capacity of a communication channel. This dissertation builds on the information-theoretic perspective to understand the limits on the ability of controllers to actively dissipate uncertainty in systems.

High-performance control systems include two types of uncertainty: noise that might be introduced by nature, and inaccuracy that might be introduced by modeling, sampling errors or clock jitter. The first of these is often modeled as an additive uncertainty and is the object of most prior work at the intersection of communication and control. This dissertation focuses on the multiplicative uncertainty that comes from modeling and sampling inaccuracies. Multiplicative uncertainty could be introduced from the observation mechanism that senses the state (the sensing channel) or from the actuation mechanism that implements actions (the actuation channel). This dissertation examines the control capacity of systems where the parameters of these channels are changing so fast that they cannot be perfectly tracked at the timescale at which control actions must be performed. This parallels the fast-fading models used for wireless communication channels.

This dissertation defines a notion of the "control capacity" of an unreliable actuation channel as the fundamental limit on the ability of a controller to stabilize a system over that channel. Our definition builds from the understanding of communication capacity as defined by Shannon. The strictest sense of control capacity, zero-error control capacity, emulates the worst-case sense of performance that the robust control paradigm captures. The weakest sense of control capacity, which we call "Shannon" control capacity, focuses on the typical behavior of the zeroth-moment or the log of the state. In between these two there exists a

range of η-th moment capacities of the actuation channel. These different notions of control capacity characterize the impact of large deviations events on the system. They also provide a partial generalization of the classic uncertainty threshold principle in control to senses of stability that go beyond the mean-squared sense.

Since the "Shannon" control capacity of an actuation channel relates to physically stabilizing the system, it can be different from the Shannon communication capacity of the associated communication channel. For the case of actuation channels with i.i.d. randomness, we provide a computable single-letter expression for the control capacity. Our formulation for control capacity also allows for explicit characterization of the value of side information in systems. We illustrate this using simple scalar and vector examples.

These ideas also extend to systems with unreliable sensing channels. Somewhat surprisingly, we find that in the case of non-coherent sensing channels, the separation paradigm that is often applied to understand control problems with communication constraints on the observation side, can fail. Active learning and control of a system state is possible (i.e. the control capacity is finite), even though passive estimation is not. An interesting aspect of this result is the techniques used. The problem is reduced to estimating the initial state of the system using a prior and the non-coherent observations. Then a genie-based argument can be used to lower bound the Bayesian problem using a minimax-style hypothesis testing bound.

The results throughout are motivated by observations made using simplified bit-level "carry-free models." These models are generalized versions of deterministic bit-level models that are commonly used in wireless network information theory to capture the interaction of signals. Here, we modify them to capture the information bottlenecks introduced by parameter uncertainty in active systems.

To all my teachers over the years, starting with my parents.

Contents

List of Figures

List of Tables

Acknowledgments

Throughout my time at Berkeley, I have had the pleasure of interacting with with many people who have had a positive impact on my life. These interactions have been among the most rewarding parts of my graduate school experience. I would like to thank all these amazing people for their help, support and friendship along the way.

First and foremost, I would like to thank my advisor Prof. Anant Sahai. I will always be grateful for his ever positive attitude, his enthusiasm towards new ideas and patience to try these ideas out. Anant let me explore a variety of problems across different areas, and showed me how to approach research. He taught me what it means to have an aesthetic, and how to have research-fun! I appreciate that he also encouraged me to help with CS70 and pushed me to become a better teacher and mentor. I feel extremely fortunate to have had the wonderful experience being his student, and I will not be able to list all the things I have learned from him here. Thank you, Anant, for your time, energy, kindness, constant support, and unending supply of band-aids.

I am also very grateful for the support and feedback from my dissertation committee. I would like to thank to Prof. Aditya Guntuboyina for discussions about bounding techniques that influenced results in Chapter 4, and Prof. Pravin Varaiya for adding his perspective to the results and pointing me to useful references. I would like to thank Prof. David Tse for his careful feedback on Chapter 5. Ideas from his class with Anant on deterministic models have heavily influenced this dissertation.

Thanks to Se Yong Park and Govind Ramnarayan, who have been my collaborators for some of the results in this dissertation. I have learned a lot from Se Yong — he has a remarkable ability to provide clear insight into the heart of any problem.

More generally, I am very grateful for having had the opportunity to work with all of my collaborators: Prof. Anant Sahai, Kristen Woyach, Kate Harrison, Se Yong Park, Pulkit Grover, Govind Ramnarayan, Vasuki Narasimha Swamy, Sahaana Suri, Matt Weiner, Paul Rigge and Prof. Bora Nikolić.

Discussions with Prof. Madhu Sudan when I was an intern with him at Microsoft Research planted the seed for some of the ideas in this dissertation. I also learned a lot working on problems in crowdsourcing with Lav Varshney, which started when I was an intern with him at IBM Watson.

Many thanks also go to the people who first introduced me to research at MIT and encouraged me to apply to graduate school: Prof. George Verghese, Tushar Parlikar, Thomas Heldt, Faisal Kashif. Similarly, I would like to thank my mentors at MIT: Prof. Denny Freeman, Prof. Collin Stultz, Prof. Dan Ariely who provided support and encouragement through my undergraduate years. I really appreciate everything they have done for me.

Members of the BMI lab — Prof. Jose Carmena, Subbu Venkatraman, Rodolphe Heliot, John Long — were my first introduction to Berkeley. Thank you all for your support during those initial years.

Many thanks go to Sheila Humphreys for pushing me, and for all her advice and guidance.

A special thank you goes out to Kristen Woyach. She has always been there for me, with both technical and non-technical support, and has (almost) always had the right answer :-). You were sorely missed this past year. Thanks also to Pulkit Grover for introducing me to Witsenhausen's counterexample. I learned a lot from our discussions.

I've had the chance to develop many "research-plus" friendships over the past few years. I'd like to thank Krish Eswaran, Baris Nakiboglu, Abolfazl Motahari, Mohammad Maddah-Ali, Giulia Fanti, Vijay Kamble, Kristen Woyach, Pulkit Grover, Kate Harrison, and Prasad Santhanam for these. I have grown a lot from our conversations.

I am very grateful to Kristen Woyach, Giulia Fanti, Nihar Shah, Varun Jog, Vasuki Narasimha Swamy and Sahaana Suri for carefully reading over drafts of this dissertation.

Many thanks go to the EECS staff members who have always gotten me out of trouble, made my life so much easier, and have become my friends over the years. Thank you to Shirley Salanio, Rebecca Miller, Ruth Gjerde, Gabe Chang, Chuck Brien, Heather Levien, Jean Richter and Kim Kail.

I would also like to thank past members of WiFo, and present members of BLISS. I am honored to be part of this group. I have benefitted from the wisdom of many more senior students and postdocs, and the camaraderie and support of my peers and junior students. Thanks to Krish Eswaran, Rahul Tandra, Mubaraq Mishra, Galen Reeves, Jiening Zhan, Hari Palaiyanur, Kristen Woyach, Pulkit Grover, Salim El Rouayheb, I-Hsiang Wang, Se Yong Park, Naveen Goela, Nebojsa Milosavljevic, Sreeram Kannan, Kate Harrison, Nihar Shah, Giulia Fanti, Vijay Kamble, Nihar Shah, Varun Jog, Vasuki Narasimha Swamy, KV Rashmi, Po-ling Loh, Ramtin Pedersani, Kangwook Lee, Sudeep Kamath, Venky Ekambaram, and Sameer Pawar.

Thank you to many other faculty in WiFo, in EECS and at Berkeley generally: I have learned a lot through great courses and other interactions with them.

One of the highlights of my graduate school experience has been working with and mentoring Govind Ramnarayan, Sahaana Suri, Ena Hariyoshi and Leah Dickstein. I have gained a lot from my interactions with them. Thank you.

Thanks to all my co-conspirators in the EEGSA. In particular, it was a pleasure working with Zach Jacobson and Matt Spencer. Thank you!

A special thanks goes to all my "first-year" friends for being there from the start, and also to many other wonderful folks in EECS and at Berkeley I have met along the way. You've made coming to Cory every morning a much more pleasant experience than it would have otherwise been! Thank you Ma'ayan Bresler, Suman Ravuri, Erin Summers, Ricardo Garcia, Lakshmi Jagannathan, Samarth Bhargava, Milos Jorgovanovic, Ankur Mehta, Claire Baek, Neeraja Yadwadkar, Mobin Javed, Floraine Berthouzoz, Kurtis Heimerl, Matt Spencer, Zach Jacobson, Maraym Kamgarpour, Achintya Madduri, Roshni Kasad, Divya Ramchandran, Shirali Pandya, Marie Lilleborge, Rikky Muller, and Anand Bhaskar. Many other friends in Berkeley and the Bay area have also been a source of joy in my life — thanks to all of them as well.

Many thanks go to my dance teacher, Guru Vishnu Tattva Das, and fellow dancers Maureen Mitra, Smitha Seshadri, and Aparna Krishnamoorthy for providing a valuable distraction from work over the years. Odissi, and all of you, mean a lot to me.

I would also like to thank Prof. Tara Javidi and Prof. Lalitha Sankar for their encouragement and support in the past few years.

My friends Mukta Gore, Onkar Dalal, Varun Kanade, Shubhangi Saraf, Sukhada Fadnavis, Shweta Shah, Gayatri Natu, Mandar Gadre, Matt Fishburn and Jennifer Gagner have provided tremendous local and remote support during the hardest times.

I would also like to thank Prakash Sir and the rest of the "Bhaskar-gang" for kindling my initial interest in Mathematics.

Last, but not least, I would like to thank my parents Neelima and Vishnu Ranade, my brother Vinayak, and my grandparents, Aaji, Babuji and Mothiaai, for their unconditional love, support and encouragement.

I gratefully acknowledge the NSF for a Graduate Research Fellowship, as well as for the grants CNS-0932410, NSF- 07291222, NSF-0917212, CNS-403427, CNS-1321155, and ECCS-1343398 that funded this work.

This dissertation is dedicated to all my teachers, starting with my parents and family. Thank you for everything.

Chapter 1

Introduction

The past few decades have seen technology and society take ever larger steps towards automation of activities in our daily lives. Today, the Internet of Things presents a vision that is strikingly close to what was futuristic science fiction fifty years ago. There has been extensive development in sensing and data processing for applications in home automation, civic and environmental systems, and medical monitoring technology. The larger vision of the Internet of Things (IoT) includes not only sensing but also simultaneous actuation of numerous wirelessly connected devices. The goal is to enable a large number of globally distributed computing devices that communicate with each other and *interact with the physical world* at both personal and industrial scale. Truly immersive IoT applications with humans in the loop require highly-reliable communication, and latency requirements on the control loop are on the order of human-response times, i.e. tens of milliseconds. Industrial control applications have similar latency constraints and require reliability on the order of 10^{-8}. The next generation of 5G wireless communication protocols are being developed today to target these applications in the high-performance regime. The processing speedup from Moore's law and more bandwidth can move us closer to this vision, but they cannot get us all the way there.

This push towards high-performance control really forces us to ask the question: what are the fundamental bottlenecks in control applications? This dissertation provides an information-theoretic approach to understand the fundamental performance limits on active control systems.

There are many parallels between the sister fields of communication and control. While communication aims to reduce uncertainty by communicating information about the state of the world, control aims to reduce uncertainty by moving the state of the world to a fixed point. Furthermore, transmitters must communicate over the unreliability in the communication channel, while controllers must overcome unreliability to stabilize the system. Extensive work in information theory has provided a framework with which we can understand the fundamental limits in communication. Communication capacity is a standard metric in wireless system design that has motivated work in coding theory and influenced protocol design in wireless communications.

This dissertation develops a notion of "control capacity" to understand the impact of uncertain parameters in control systems. High-performance systems might have systems parameters that cannot be perfectly tracked at the timescale at which control actions must be performed. Further, systems can include both uncertainties that are introduced by nature but also inaccuracies that are introduced by the system model itself. While nature-introduced uncertainties tend to be additive, model parameter errors can introduce multiplicative uncertainties in the system.

Within control theory, system identification and adaptive control have developed algorithms by which a controller can gather information to better model a system. The controller first uses known perturbations to learn the system dynamics to a desired accuracy and then stabilizes it. This may be done in an active fashion where the input signals are chosen online, in response to the system's observed behavior. Alternatively, it might be done passively with a predetermined set of inputs. Robust control tends to deal with uncertainty using a non-stochastic worst-case perspective, and often focuses on linear strategies for stabilization.

However, traditional information theory ignores message semantics and focuses on system rate. Complexity, delay and reliability are secondary considerations. On the other hand, traditional control theory has a natural focus on interpreting an observation to perform the right action and cares about the rate of convergence and computable solutions. As a result the tools in both of these fields have largely been developed independently of each other and do not quite mesh. There have been previous attempts to try to develop a theory of information that is compatible with the cost function and structure of control systems, as indicated by this quote from Witsenhausen in 1971 [114]:

> Changes in information produce changes in optimal cost. This suggests the idea of measuring information by its effect upon the optimal cost, as has been proposed many times [11]. Such a measure of information is entirely dependent upon the problem at hand and is clearly not additive. The only general property that it is known to possess is that additional information, if available free of charge, can never do harm though it may well be useless.This simple monotonicity property is in sharp contrast with the elaborate results of information transmission theory. The latter deals with an essentially simpler problem, because the transmission of the information is considered independently of its use, long periods of use of a transmission channel are assumed, and delays are ignored.
>
> Efforts to establish a new theory of information, taking optimal cost into account, have not as yet been convincing [54, 104, 113, 8].

This dissertation proposes a novel information-theoretic approach to understand the bottlenecks due to parameter uncertainties in control systems. The notion of "*control capacity*" proposed depends on the problem formulation and is also additive, i.e. it can easily account for additional side information.

1.1 Basic system model

To set up a framework for the dissertation, we give the general system model (Fig. 1.1). Elements on the observation and control side mirror each other.

This thesis considers three types of channels in the system: communication channels, sensing channels, and actuation channels. Communication channels are the channels that are traditionally considered in information theory, and may occur on the input or on the output side of the plant. These model unreliability and noise introduced after the data collection process is completed or before the actuation has started. As a result, encoders and decoders can be used to combat the uncertainty of these channels. The remaining two are illustrated in the figure below. The bottlenecks presented by the sensing and the actuation channels are related to the physical limitations of the devices.

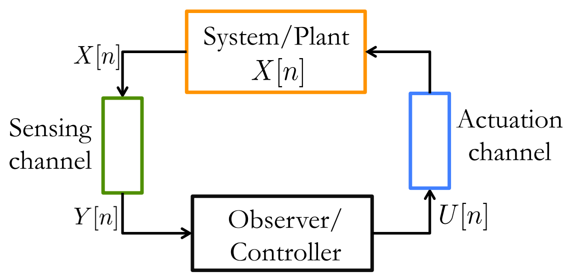


Figure 1.1: A general model for control and estimation of a system over networks.

- System/plant: The system determines the physical evolution of the system state.
- Observer/controller: This is a box with computational power that can estimate the system state and compute controls. The desired control actions are sent to the actuator via the actuation channel. There might be communication channels inside this box that use encoding and decoding.
- Sensors and sensing channels: The sensors and sensing channel is the output interface of the system to the rest of the world. The sensing channel has access to (noisy) measurements of the system state, but has little to no computational power. It transmits the observations out of the system. In a sense, decoding of these transmissions by the controller is possible, but encoding isn't. This might also be a network. We can think of this as the physical sensor that is actually collecting the data.
- Actuators and actuation channels: The actuation channel is the input interface for the controller to interact with the system. Control signals are transmitted over this to the system actuators. This also has little to no computational power and can only execute the controls sent to it by the controller. This represents the physical actuator that is implementing the control. So here, encoding might be possible, but decoding isn't.

Standard linear system models for estimation often have the format below:

$$\begin{aligned} \vec{X}[n+1] &= A \cdot \vec{X}[n] + \vec{W}[n] \\ \vec{Y}[n] &= C \cdot \vec{X}[n] + \vec{V}[n] \end{aligned} \tag{1.1}$$

$\vec{X}[n]$ is the $m \times 1$ state vector and A is a known $m \times m$ constant matrix. n is the time index. C is also known and may or may not be full rank. $\vec{W}[n], \vec{V}[n]$ are a additive $m \times 1$ noise matrices. With no observation encoding or rate-limits, the optimal open-loop estimation in the presence of additive noise is given by the traditional Kalman filter [58].

This setup is slightly modified in the case of control, with the control signal $\vec{U}[n]$ which is a function of $\vec{Y}[n]$, added in. B is the control gain matrix.

$$\begin{aligned} \vec{X}[n+1] &= A \cdot \vec{X}[n] + B \cdot \vec{U}[n] + \vec{W}[n] \\ \vec{Y}[n] &= C \cdot \vec{X}[n] + \vec{V}[n] \end{aligned} \tag{1.2}$$

The objective of a good controller is to minimize the cost function associated with the system. This cost usually involves both the state magnitude and the control power. Papers such as [101] and related works successfully use dynamic programming to tackle these problems. However, it can be difficult to deal with rate-constraints in a system in the presence of a control cost. To simplify the setting, many works in the community tend to focus on asymptotic system stability, since it is the precursor to understanding the full control cost. This dissertation also focuses on stability as a first step to understand control capacity.

Much of the previous work at the intersection of communication and control has focused on the channels on the output of the system, i.e. the observation and sensing side. This is because there is an obvious connection between communication and control when we think of the state of the control system as an information *source*. The system is generating messages that must be transmitted to the controller at adequate rate so that the system can be tracked and also stabilized. In this context, estimation theory is the natural bridge from communication to control.

The next section gives a history of related work at the intersection of communication and control.

1.2 Previous work

We divide the related previous work into four subsections. The first subsection focuses on performance bottlenecks in the *sensing and observation channels* in a control system, while the second focuses on performance bottlenecks due to the *control and actuation channels* in the system. When thinking of communication (or sensing or actuation) channels that are being used for control, the required rate and required reliability are the two relevant measures. We discuss both of these here. The impact of observation-side bottlenecks are much better understood than the control side bottlenecks, and this thesis tries to fill in some of the gaps on the latter side. There has been very little work on actuation channels that is compatible with information-theoretic tools. Many of the arguments are separation-based.

Two survey papers give a very good idea of the state-of-the-art for both these subsections. The survey paper by Schenato et al. [101] highlights the use of Riccati equations and LMI techniques to provide numerically computable bounds and thresholds for system stability. A second survey paper (from the same IEEE special issue) by Nair et al. [80] focuses on rate constraints in control systems and uses statistical and information-theoretic tools. The models in [101] use an infinite capacity real-erasure channel in their model. The formulations considered in [101] do not allow for encoding and decoding over the sensing and observation channels. This allows for a focus on the the reliability bottleneck that is generated by observation and sensing channels: errors are caused when information is received "too late" as opposed to because "too few" bits are received. On the other hand, the models in [80] are rate-constrained and the results extensively use information-theoretic tools. However, these models cannot capture the critical aspects of reliability for control. The work discussed in [80] considers powerful encoders and decoders on the sensing and observation channel in the system. The following sections discuss the relationships between both sets of models as well as the complementary perspectives they provide.

The third subsection provides a discussion of uncertain parameters and *multiplicative noise* in systems, which is the focus of this dissertation. Finally, we discuss previous discussions on the *value of information* in systems.

1.2.1 Sensing and observation channels

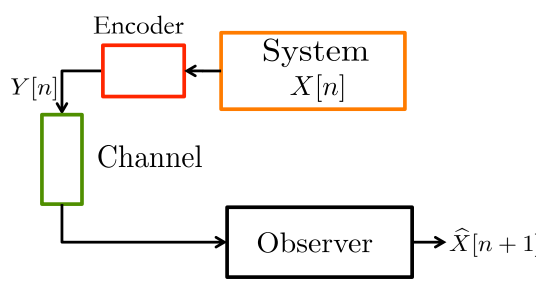


Figure 1.2: Estimation of the system state based on observation-feedback over a lossy link. The decoder is part of the observation box.

Information theorists look to rate as the parameter that characterizes a communication system. A natural first attack is to bring in tools from information theory and look at estimation and control over rate-limited channels.

Rate-limited observation channels

Explicit encoding and decoding with rate-limited channels was was first considered in papers such as Wong et al. [117], Nair et al. [79] and Tatikonda and Mitter [110]. [110] looked at the estimation and then control of an unstable system over a noiseless rate-limited channel, with access to a full encoder and decoder on either end of the channel (Fig. 1.2). They showed there is a threshold on the rate required over a communication channel for stable estimation and stable control, and that the threshold is the same in both cases! First, they showed that for an observer to be able to track the state with finite mean squared error, the observation channel must be able to support a rate $R > \sum \log |\lambda(A)|$, where $\lambda(A)$ are the unstable eigenvalues of the matrix A. The proof technique was non-standard in control at the time, and hinged on showing that the volume of uncertainty in the system was growing as the product of the unstable eigenvalues of A.

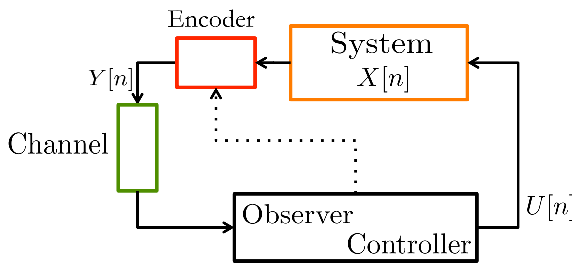


Figure 1.3: Rate-limited control allows an encoder and decoder for communication across the channel. The decoder box can be thought to be merged with the observer/controller box and is not shown separately here. Further, the controller provides feedback of the control signal sent to the encoder box so that only the innovation in the system need be encoded.

Their basic argument to show that the threshold for stabilization is the same as for estimation is based on the separation paradigm. First, they note that the controller can simulate an observer and estimate the system state in open-loop. Since it has memory of the previous controls it applied, it can also estimate the closed-loop system state. Then, it can compute the optimal control from the current state-estimate as the one that will cancel out the state. The necessary condition on the rate is given from the estimation results. Subtracting the state estimate at each time (using knowledge of the controls that have been previously applied by the controller) gives a matching sufficient condition. Full knowledge of the actual impact of the applied controls is critical in this strategy, and the results in [110, 96] use control feedback to the encoder to achieve stabilization (see Fig. 1.3).

A large set of data-rate theorems for control over (noiseless and noisy) rate-limited channels have built on these ideas and follow from separation theorems, as in [110, 96, 79], and more recently in results like [75]. In these problems, it is common to think of the observer and controller as lumped into one box (Fig. 1.3), with a noiseless channel between controller and actuator.

One property of messages/bits transmitted in a communication system is the transmission rate. A second one is the reliability at which the bits are received by the decoder. This reliability is particularly important in control systems [97]. However, this reliability, or the probability of error associated with a bit is ignored in noiseless rate-limited channels. [97] used a binary erasure channel to show that just large enough Shannon capacity is not sufficient to ensure bounded estimation error. The most significant bits (MSBs) of the state must be estimated to greater reliability, i.e. the communication system must guarantee a lower probability of error for the MSBs of the state as time goes on. This was fully understood for scalar systems by Sahai and Mitter [96]. This paper showed that for a scalar system with eigenvalue a, a necessary and sufficient condition for a finite η-moment error is that the anytime capacity for reliability $\eta \log a$ is greater than $\log a$.

Using the discussions in [98] and deterministic style models such as those proposed by Avestimehr, Diggavi and Tse (ADT models), we can illustrate this result. Think of the system state as a stack of bits, with the LSBs at the bottom and MSBs at the top. This stack is growing taller with each clock tick as the unstable system is multiplied by a gain (Fig. 1.4). The observer must estimate each bit before it gets too large, and so the required reliability on the bits increases as they take higher positions in the stack. Since the cost bits are proceeding upwards by $\eta \log a$ positions each time, i.e. the magnitude of the η-moment of the state is growing by $2^{\eta \log a}$, the probability of error on that state bit must shrink by at least $2^{-\eta \log a}$. Hence, $\eta \log a$ is the required reliability. The communication rate is given by the number of new bits that are pulled into the system state at each time, which is $\log a$. (We emphasize this picture because it has been useful to understand the information flow in other decentralized estimation and control problems [47, 84, 93]).

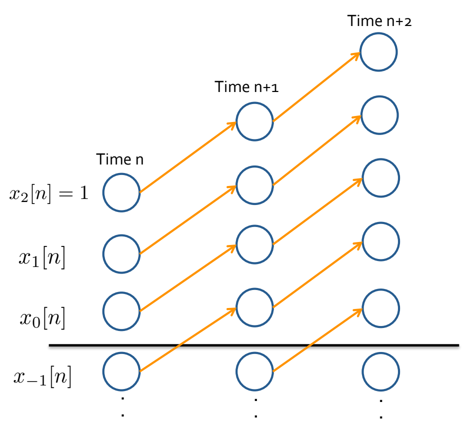


Figure 1.4: The system state can be thought of as a stack of bits marching upward.

The scalar anytime ideas can be extended to vector systems through a change of coordinates and diagonalization [99]. Projected on to the appropriate basis, a vector system is like multiple (possibly interacting) scalar systems. This may come at the cost of finite delay: if the observations belong to a lower dimensional subspace than the system due to the structure of the B matrix, then the system must wait to accumulate enough observations. [99] gives a characterization of stable regions involving reliability vectors. While we can compute these in the scalar case, for the vector setup the regions are hard to compute even for simple channels like the BEC with feedback. We do not even have non-trivial bounds. Hence, the result is somewhat dissatisfying.

The special case of the real-erasure channel (with arbitrary encoding and decoding) is easier to understand [99]. The rate constraint is removed with the infinite-capacity-real-erasure channel, and the limit arises due to the reliability constraint alone. Then, the critical erasure probability is dictated only by the maximal eigenvalue of the A matrix, and is $\frac{1}{\lambda_{max}(A)^{\eta}}$ for η-moment stability of the error[1].

This shows that reliability constraints have a "max" nature: they combine to give the strictest of the constraints (when rate is not an issue). On the other hand, rate constraints naturally add, as we know from Tatikonda and Mitter [110].

Real-erasure observation channels

The vector real-erasure channel in the anytime results requires a very non-trivial encoding to be able to meet the $\frac{1}{\lambda_{max}(A)^{\eta}}$ constraint: a vector is packed into a scalar by bit interleaving! Does the result hinge on this complex encoding strategy?

This question leads to the intermittent Kalman Filtering model, which also considers mean-squared estimation of a system over a real-erasure channel [102]. This model simply takes encoding off the table and focuses on reliability. A basic block diagram is in Fig. 1.5.

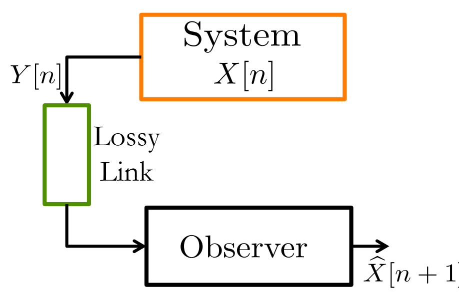


Figure 1.5: Estimation of the system state based on observation-feedback over a lossy link. Note that encoding is not allowed.

[1]Consider the following strategy: at each time encode the real-valued observations for each element of the system state vector into one real number by interleaving the bits. We have access to a powerful decoder and so the state-estimation error is renewed every time an observation arrives. As long as the erasure probability is low enough that it can catch the growth due to the maximal eigenvalue, the estimation error remains bounded. This is caricatured by the deterministic model picture of bits running off to infinity along the eigenvectors for each eigenvalue of A.

To be able to use the traditional Kalman Filtering tools, Sinopoli et al. modeled packet drops over a network as an additive noise with infinite variance, i.e. a noise that wiped out the observation entirely. Subsequently, they showed that the results also hold with packet drops modeled as a discrete-valued multiplicative noise, $\beta[n] = \text{Bernoulli}(1 - p_e)$, in addition to the additive noise for the system observation [101], as below:

$$\begin{aligned} \vec{X}[n+1] &= A \cdot \vec{X}[n] + \vec{W}[n] \\ \vec{Y}[n] &= \beta[n](\vec{X}[n] + \vec{V}[n]) \end{aligned} \tag{1.3}$$

Using a Riccati equation approach, they show that the optimal Kalman gain for this problem is time-varying and depends on the arrival sequence $\beta[n]$. They then restrict the solution to be linear and time-invariant to give a Modified Algebraic Riccati Equation which can be analyzed using LMI techniques as shown in [102]. With this, they show that the critical erasure probability for the mean-squared error is upper-bounded by $\frac{1}{\lambda_{max}(A)^2}$, where $\lambda_{max}(A)$ is the maximal eigenvalue of the matrix A. The estimation error is unstable for higher packet loss probabilities. This matches the result of the vector anytime problem with $\eta = 2$ in the previous section: a strategy without encoding cannot beat a strategy with encoding. However, this bound is not tight, and they find that in certain cases the LMI predictions do not match this bound. This left a gap open to explore.

In a decentralized setup, simultaneous encoding of the different components of a vector system-state might not even be possible. What happens if all of the required observations are not generated by the same sensor and hence have different erasure probabilities? This question was noted and explored by Liu and Goldsmith [65]. By holding one of the two drop probabilities constant, they formulate a problem similar to [102] and provide LMI-based bounds for the critical erasure probabilities. In this case as well, there is gap between the upper and lower bounds in the general case. However, the authors show how the bounds can help design decisions. Engineering considerations might allow a designer to allocate more or less resources to a particular channel and these bounds can guide the allocations to ensure system stability.

Mo and Sinopoli [77] further explored the gap above and gave clear examples of systems where the critical erasure probability can be either $\frac{1}{\lambda_{max}(A)^2}$ or $\frac{1}{\Pi\lambda(A)^2}$ based on the nature of the matrix A with no restriction on the filter gain (i.e. it was allowed to vary with time and depend on the specific observation arrival sequence). This gap was then fully understood by the introduction and analysis of eigenvalue cycles in [86]. If the pattern of observation erasures is such that the observer keeps receiving redundant information, it will not be able to decode the state. It turns out that this can only happen in special cases where the eigenvalues of A have a very specific cyclic structure. For instance, if the system dimension and cycle period are both n, then the critical erasure probability is $\frac{1}{\Pi\lambda(A)^2}$. However, for generic A the critical erasure probability is exactly $\frac{1}{\lambda_{max}(A)^2}$, cleanly resolving the estimation problem [86]. In principle, these results have brought closure to the erased observations problem.

It is important to note that the works above did not allow for the possibility of an encoder with memory or delay. The estimate for $X[n]$ must be produced at time n. One of the ideas in [86] is to break up the eigenvalue cycles by non-uniform sampling. This could also be achieved by an encoder of memory-n that takes a random linear combination of samples before sending them across the channel, which would get us back to the $\frac{1}{\lambda_{max}(A)^2}$ bound [86].

1.2.2 Control channel

In addition to unreliable and noisy sensing and communication channels on the output side, real-world systems also have to contend with unreliable or delay-prone actuation channels or communication channels on the control end.

Consider

$$\vec{X}[n+1] = A \cdot \vec{X}[n] + \gamma[n] \cdot B \cdot U[n],$$
$$\vec{Y}[n] = \vec{X}[n], \tag{1.4}$$

where $\vec{X}[n]$ is the $m \times 1$ state vector, A is an $m \times m$ constant matrix, B is an $m \times 1$ constant matrix, and $\gamma[n]$ are i.i.d. Bernoulli$(1 - p_e)$ random variables.

Elia and co-authors were among the first to consider control actions sent over a real-erasure actuation channel (eq. 1.4)and ask: how can we control a system (with perfect observations) when the control packets to the actuator might be lost over a real-erasure channel [32, 30, 31]? They restricted the search space to consider only linear time-invariant (LTI) strategies so that the problem might become tractable[2]. [101, 55] also looked at this problem using dynamic programming techniques and showed that LTI strategies are in fact optimal in the infinite horizon. Riccati equation techniques and dynamic programming can show a separation result in the case of simultaneous lossy observation and control channels [101, 55].

[32, 101, 55] also identified the importance of acknowledgements on the actuation channel. In the case where packet losses over the actuation channel are acknowledged (i.e. the control transmission is over a "TCP-like" protocol, Fig. 1.6), separation between estimation and control holds. The optimal control strategy is a linear function of the state-estimate, and the optimal controller gain converges to a constant in the infinite horizon case [101]. On the other hand, in the case where control packets are not acknowledged (i.e. the actuation transmission uses a "UDP-like" protocol), separation does not hold. In the generic case, non-linear and possibly time-varying strategies are necessary [101, 55]. The estimator error

[2]This restriction is essentially the same as in [102]. Elia showed that $p_e < \frac{1}{\Pi \lambda(A)^2}$, where $\lambda(A)$ are the unstable eigenvalues of the matrix A, is the critical probability for stabilization using LTI controllers. The techniques here provides an insight into the theorem that is hidden when seen through the dynamic programming lens and are discussed further in Chapter 7.

Sinopoli et al. restricted the Kalman gain to be time-invariant to be able to approach the problem using standard Riccati techniques.

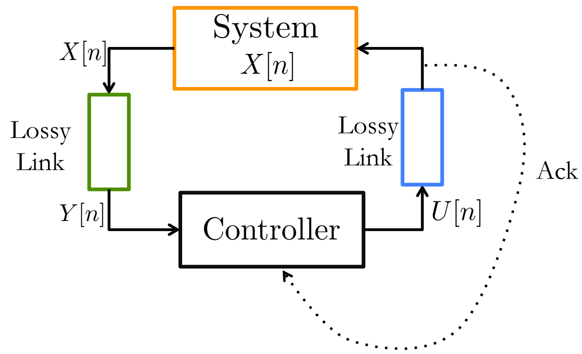


Figure 1.6: Control packets are acknowledged as in a TCP protocol.

covariance depends on the control input, and the erasure probabilities of the observation and the control channel may interact.

Combining the fact that LTI strategies are optimal (with packet acknowledgements) [101, 55], with the bound for LTI controllers from [30] we can conclude that the $p_e < \frac{1}{\Pi\lambda(A)^2}$ bound holds for all control strategies[3]. The restriction to LTI strategies in [30] was in fact not a restriction at all!

Liu and Goldsmith [65] considered the case of physically separate sensors for estimation. What about the control problem with physically separate actuators and separate channels between different actuators and the controller? This problem is considered by Garone, Sinopoli, Goldsmith and Casavola in [36, 37]. They extend the techniques of [101] to show that just as with a single channel, the separation principle still holds with both multiple sensor and actuator channels and acknowledgements for the control channel. In [37], it is interesting that the optimal control is a linear function of the state that explicitly depends on the loss probabilities of each channel (Remark 4 of [37]). This is different from the case of a single channel, where the controller can afford to remain agnostic to the drop probabilities. There, the optimal control at time k is computed assuming that it will get through the channel.

We provide a simple scalar example (not in the paper [37]) helps illustrate what is going on here:

$$\begin{aligned} X[n+1] &= A \cdot X[n] + 2 \cdot \gamma_1[n] \cdot U_1[n] + \frac{1}{2} \cdot \gamma_2[n] \cdot U_2[n], \\ Y[n] &= X[n] \end{aligned} \tag{1.5}$$

Let $\gamma_1[n]$ and $\gamma_2[n]$ be independent Bernoulli-(p_1) and Bernoulli-(p_2) random variables. Then, the optimal control vectors $U_1[n]$ and $U_2[n]$ must take p_1 and p_2 into account: for instance the control with $p_1 = 0$ and $p_2 = 1$ is clearly different from $p_1 = 1$ and $p_2 = 0$ and also

[3]The assumption of noiseless perfect observations in [30] is equivalent to assuming control acknowledgements.

different from $p_1 = p_2 = 0.5$. In the last case, some kind of average control action is the right thing to do. The two independent random variables combine to form a random multiplicative scaling for the control action. The control must account for this multiplicative noise and hence depends on the drop probabilities.

Gupta, Spanos, Hassibi and Murray [52] used dynamic programming techniques to show that an observation/sensing channel can tolerate more erasures if the system is allowed an encoder and decoder around the channel: the critical probability goes from depending on the product of the eigenvalues to depending only on the maximal eigenvalue. While this can be seen directly using an information-theoretic perspective, it is more challenging to arrive at the result using other arguments. The problem considered by Gupta, Sinopoli, Adlakha, Goldsmith and Murray in [51] builds on this. Here authors consider the utility of an actuator with memory, which can be thought of as an approach to understand the control case with encoding over control packets allowed. For a generic A matrix with no cycles and rank-1 B matrix, an actuator memory equal to the dimension of the system changes the tolerable erasure probability from $\frac{1}{\Pi\lambda(A)^2}$ as in [30] to $\frac{1}{\lambda_{max}(A)^2}$. We provide a simple example (not in the paper [51]) to understand the spirit of the result:

$$
\begin{aligned}
X[n+1] &= \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} X[n] + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} U[n], \\
Y[n] &= X[n],
\end{aligned} \tag{1.6}
$$

with distinct $\lambda_1, \lambda_2, \lambda_3$. Then the system can be set to 0 in three time steps. If these controls can be precomputed and sent to the actuator in advance (i.e. the actuator has a memory buffer of 3), one time step is sufficient, which brings us to the $\frac{1}{\lambda_{max}(A)^2}$ bound since now we only have to account for the fastest eigenvalue. This helps interpret the simulation results presented in [51]: for a two dimensional system the control cost became constant with actuator memory of 2. The authors use these simulations to point out that such memory would be most useful in the case of correlated or bursty losses, which are common in real-world networks.

Unlike the estimation setup, which is now well understood, the space of control problems still has open questions. So it seems that the duality between estimation and control does not always hold. Consider the simple case of a scalar observation sent over a real-erasure channel for estimation and the dual control problem with perfect observations but a scalar real-erasure control channel. We know from [86] that the critical drop probability for estimation of a system, as in eq. (1.3), with generic matrix A is $\frac{1}{\lambda_{max}(A)^2}$. This is different from the erasure probability tolerable in the dual control problem, as in eq. (1.4), considered by Elia [30] which is $\frac{1}{\Pi\lambda(A)^2}$. This is significantly more stringent than in the observation case.

Why does the duality of estimation and control break down in the case of lossy sensing and actuation channels? Why is the control problem harder? We will try to address these questions in this dissertation.

1.2.3 Parameter uncertainty as multiplicative noise

Parameter uncertainty has a long history in controls, communication and learning, and usually manifests as multiplicative randomness in the system. The erasure sensing and actuation channels considered in the earlier sections are specific examples of multiplicative uncertainty on a channel, Bernoulli uncertainty. This dissertation explores more general multiplicative noise on the sensing and actuation channels in a system.

Since linear systems are additive and commutative, additive noise in the actuation channel has the same effect as additive noise in the plant itself would. Hence, the data-rate theorems discussed in the earlier section effectively capture the effect of additive uncertainty on the actuation channel. However, the examples in the previous section show that multiplicative noise can break this duality between the sensing and actuation channels, and we explore the two aspects separately in this thesis.

Communication systems

Communication channels with multiplicative noise have been extensively studied in information theory. Non-coherent channels in communication systems emerge when phase-noise, frequency hopping or fast-fading make it impossible for the receiver to perfectly track the channel state. Such communication has been studied as far back as 1969 when Richters conjectured that even for continuous unknown fading-distributions the optimal input distribution is discrete [94]. The conjecture was proved in 2001 [1, 59]. Since then we know that the capacity of channels with unknown Gaussian fading scales as $\log\log(SNR)$ as opposed to the $\log(SNR)$ scaling of channels with known fading [108, 63].

Control and estimation systems

Non-coherent observations in control systems can similarly arise from synchronization and sampling errors in control systems [73]. Consider a nearly trivial continuous-time system:

$$\dot{X}(t) = a \cdot X(t).$$

The system collects samples at regular intervals $t_0, 2t_0, 3t_0$ and so on. The system differential equation above implies $X(t) = e^{at}X(0)$, which gives the difference equation:

$$X[n+1] = e^{at_0} \cdot X[n].$$

Say we are recording a sample $Y[n]$ at time nt_0, but clock noise leads to the sample actually being collected at time $nt_0 + \Delta t_0$, where Δt_0 is a continuous-valued random noise variable. Then we have:

$$Y[n] = e^{a(nt_0+\Delta t_0)}X(0) = \beta[n]e^{a(nt_0)}X(0).$$

where $\beta[n] = e^{a\Delta t_0}$ is necessarily a continuous-valued random variable. We would have liked to observe $Y[n] = e^{a(nt_0)}X(0)$, but we must make do with the multiplicative random noise. Thus, with timing jitter, state information is received over a non-coherent channel.

Estimation with multiplicative noise was explored around the same time as questions regarding non-coherent transmission were being asked [91, 112]. Furthermore, the very idea of robust control [122] and adaptive control [5] is about dealing with parameter uncertainty.

On the control side, the paper [29], by El Ghaoui is an example of a discussion of control of systems with multiplicative noise. This paper takes a convex optimization approach and solutions involve linear matrix inequalities. The author states that: *Some of these problems have a known "analytical solution" (via a nonstandard Riccati equation...); most do not.* After this, multiplicative uncertainty has also been explored in a robust control context for example by Phat et al. [89] or Gershon et al. [41]. These papers take a worst-case perspective and primarily use LMI techniques. There is also work in robust control that focuses on the $\mathcal{H}_2$ norm and takes a stochastic approach.

Elia [30] builds on the robust control approach. Using a restriction to LTI strategies, he is able to connect the results back to the system eigenvalues. However, in general, while LMI based approaches allow us to numerically get a handle on the system parameters, they do not provide much intuition into the fundamental bottlenecks of the problem and are not compatible with information-theoretic approaches. Other attempts to characterize the behavior of systems with parameter uncertainty (i.e. multiplicative noise) have been limited to linear strategies [30, 91, 112] or *a priori* stable systems [42, 40, 121].

A classical result in control theory that deals with multiplicative uncertainty in systems is the uncertainty threshold principle [6]. They consider a scalar linear system with Gaussian randomness in both the system gain and the control gain. The observation channel is perfect. Using Riccati equation techniques as developed in Aoki [3], they provide an explicit threshold on the uncertainty in the system and control gain that can be tolerated for mean-squared stabilization of the system. Unfortunately, there is no analytical solution for the imperfect information counterpart of this model [12, p. 229]. Furthermore, these general dynamic programming techniques are difficult to apply beyond the quadratic cost case.

The separation theorem often comes to the rescue in the imperfect information case [12, p. 233], and dynamic programming techniques can give linear strategies for quadratic cost case. However, it is more difficult to understand general cost functions on the state. Furthermore, as we will see later in this thesis, the separation theorem might not hold in examples with multiplicative noise.

Finally, it is worth mentioning that parameter uncertainty and multiplicative noise has also been studied in through the reinforcement learning paradigm. Bandit problems focus on the tradeoff between *exploring* unknown parameters in a system and taking actions based to *exploit* the system [16].

1.2.4 Value of information

As mentioned by Witsenhausen [114], the idea of measuring information by it's impact on optimal cost has be proposed many times. In fact, the uncertainty threshold principle can be thought of as one way to measure this. The change in the threshold based on information that decreases the entropy of the system or control gains can be interpreted as the value of that information. Can we extend these ideas beyond the Gaussian mean-squared case?

There have also been many other efforts to extend the ideas of information theory into a domain that also takes into account the semantics of the information being communicated (the optimal cost in the the case of control problems.) Goldreich et al. [44], Juba and Sudan [57] and Sudan [105] take a computational complexity approach to the problem. They define communication to be successful if two agents can achieve a common goal as determined by an oracle, and indirectly explore notions of value of information in "universal" systems.

There have been many attempts to understand the value of information within control theory, such as the works described by Witsenhausen [54, 8]. Other later works include Davis [26] and Back and Pliska [9]. These are discussed in further detail in later chapters in the thesis. However, these all focus on a mean-squared sense of stability and are not compatible with information-theoretic notions. For example, side information is hard to understand in these contexts. This dissertation tries to generalize beyond those concepts.

1.3 Contributions

This thesis develops a framework and a set of tools to illustrate and quantify information flows in active control systems with uncertain parameters. One of the key developments is the notion of control capacity, an information-theory-compatible measure of the fundamental limit on the stabilizability of a system. This is developed towards the end of the thesis starting in Chapter 5. The presentation starts with building bit-level carry-free deterministic models for uncertain parameters in communication networks in Chapter 2. These ideas are then ported to the control and estimation framework in Chapter 3. The investigation of real-valued control and estimation systems begins with multiplicative uncertainty on the sensing channel in Chapter 4. We build on this to understand parameter uncertainty in the control channel and the value of side information in Chapters 6 and 7.

The organization of the thesis and key results in each chapter are discussed below as a guide to the reader.

1.3.1 Bit level models

Chapter 2 builds from the deterministic models for information flows in wireless network information theory presented by Avestimehr et al. [7] and Niesen and Maddah-Ali [82] to understand communication in networks with unknown fading. These models provide a natural intuition for the $\log \log SNR$ capacity of non-coherent channels, i.e. channels with unknown fading [63]. In the unknown fading context, a carry-free model can be further simplified to a max-superposition model, where signals are superposed by a nonlinear max operation. Surprisingly, unlike in relay-networks with known fading and linear superposition, we find that decode-and-forward can perform arbitrarily better than compress-and-forward in max-superposition relay networks with unknown fading. "Beam-forming" is critical in these networks despite the parameter uncertainty.

We extend carry-free models to understand control and estimation systems in Chapter 3. We prove a version of the uncertainty threshold principle in the carry-free models. Surprisingly, these models show that the underlying bottlenecks that give rise to the $\log \log SNR$ result in communication and the uncertainty threshold principle are exactly the same: they stem from the bit scrambling that happens due to multiplicative noise. These models highlight difference between multiplicative and additive uncertainties and form a basis for the discussions in the subsequent chapters.

1.3.2 Uncertainty in the sensing channel

Chapter 4 studies the problem of estimating and controlling an unstable system over a non-coherent channel. We find that carry-free models predict a result that is quite hard to predict or prove using standard techniques in estimation and control. Systems with non-coherent sensing channels do not obey the separation theorem, which is somewhat surprising in the context of the large body of work on data-rate theorems for the sensing channel. It is shown that such an unstable system is not mean-squared observable regardless of the density of the random observation gain: the mean-squared estimation error for any estimator must go to infinity. Standard dynamic programming techniques as well as estimation error lowerbounds like the Cramer-Rao bound fail to give this result. We use a novel genie based proof to convert the Bayesian problem to a minimax style hypothesis testing problem to give the lowerbound.

Achievable strategies to stabilize plants with non-coherent observations have been well known. Carry-free models help illustrate why active learning helps extract information and stabilize the system even though passive estimation of the system is not possible. This is not clear from standard dynamic programming techniques as well as information theory rate-distortion approaches.

1.3.3 Control capacity: uncertainty in the actuation channel

Chapter 5 finally delivers the promised perspective on control capacity. We provide definitions for the control capacity of carry-free models as well as for simple scalar models of uncertain actuation channels. Our definition builds from the understanding of communication capacity as defined by Shannon. The strictest sense of control capacity, zero-error control capacity, emulates the worst-case sense of performance that the robust control paradigm captures. The weakest sense of control capacity, which we call "Shannon" control capacity, focuses on the typical behavior of the zeroth-moment or the log of the state. In between these two there exists a range of η-th moment capacities of the actuation channel. These different notions of control capacity characterize the impact of large deviations events on the system, and provide a partial generalization of the classic uncertainty threshold principle in control to sense of stability that go beyond the mean-squared sense.

Since the “Shannon” control capacity of an actuation channel relates to physically stabilizing the system, it can be different from the Shannon communication capacity of the associated communication channel. For the case of actuation channels with i.i.d. randomness we provide a single-letter expression for the control capacity. Our formulation for control capacity allows for explicit characterization of the value of side information in systems. We believe these ideas will extend to vector systems as well as to uncertainties in the sensing channels and the system gain itself, but this development remains as future work.

1.3.4 The value of information in control

Chapter 6 focuses on the value of side information in control systems from the perspective of the “tolerable growth rate” for a system. This is the maximal gain that the system can tolerate while there still exists a strategy that can potentially stabilize it. Information theory and portfolio theory have extensively studied the growth rates of systems (e.g. the size of the message set and the value of a portfolio). Control systems on the other hand care about “decay rates” since the objective is to shrink the system as fast as possible. As in portfolio theory, we can think of the value of side information in a control system as the change in the “growth rate” due to side information. A scalar counterexample (motivated by carry-free deterministic models) shows the value of side information for control does not exactly parallel the value of side information for portfolios. Mutual-information does not seem to be a bound here.

We extend this concept to vector systems through a “spinning” system toy example. The control directions for this system are re-oriented at each time so that the control or observation direction is partially unknown. The value of side information can be calculated in this setup and it behaves quite differently in a control vs. estimation context.

1.3.5 Future information in control

Chatper 7 circles back to the classically studied vector observation and control problems over (scalar) real-erasure channels studied by [102, 86] and [30, 101] respectively. In this case we try and measure the value of side information though the change in the critical packet-drop probability for the system. We show that in the case of maximal periodicity in the eigenvalues of the system matrix, the dual observation and the control problems actually will have exactly the same critical threshold probability, which is bottlenecked by the maximal eigenvalue of the system matrix.

However, we find that in the generic case, with no eigenvalue cycles, non-causal side information regarding the actuation channel drops can be very valuable for the control problem. The same side information would not improve the performance of the estimation problem. Full-lookahead on the drops by the actuation channel can help the control problem match the performance of the observation problem! Using dynamic programming, we can also show that even partial lookahead can improve the scaling behavior of the critical probability for the control problem. It seems that an align-and-kill strategy, which is similar to the zero-forcing equalization strategy in communications is the optimal strategy with and without lookahead in the control counterpart to the “high SNR” regime (i.e. as system eigenvalues tend to infinity).

Chapter 2

Bit-level models for uncertain parameters: communication

2.1 Introduction

This chapter develops a new set of bit-level channel models, called "carry-free" models for uncertain and unknown communication channels. These build on previous bit-level models developed in wireless network information theory. Chapter 3 extends these to dynamical systems.

Channel models have always formed the foundation for understanding and engineering wireless communication systems. The choice of model involves a tradeoff between simplicity and "distance" from the real-world. Fig. 2.1 shows a spectrum of these models. Rayleigh-faded Gaussian models are among the most commonly used and have undoubtedly furthered the state-of-the-art in the field. However, these models are complex, and this complexity can obscure the essential interaction between channels and the information flowing through them.

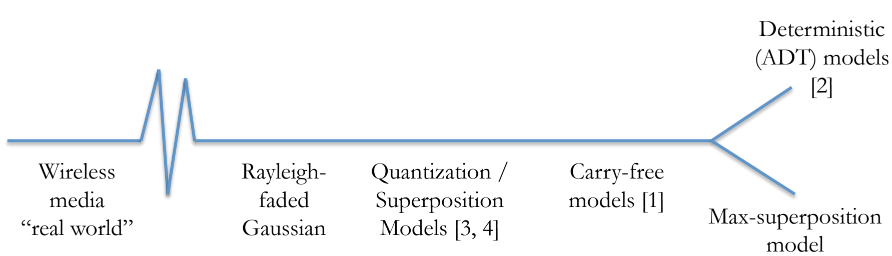


Figure 2.1: Models for wireless media. Carry-free models lie between the Rayleigh-faded Gaussian and quantization models on one side, and ADT models on the other.

This led to the development of the simple deterministic "bubble-model" by Avestimehr,

	Faded Gaussian	Quantization [2, 14]	Carry-free [82]	ADT [7]
2×2 IC	✓	✓	✓	✓
Relay channel	✓	✓	✓	✓
$(\frac{1}{2}, \frac{1}{2})$ d.o.f. in MAC w/o timesharing & CSIT	✓	✓	✓	×
Real IA	✓	✓	✓	×
Direct map to Gaussian Networks	✓	✓	×	×
Beam-forming gain	✓	✓	×	×
Unknown fading	✓	✓*	✓	×

Table 2.1: A comparison of communication models. The ✓* under Quantization models indicates that while these models can be extended to capture the effects of unknown fading in channels, the references [2, 14] do not show this.

Diggavi and Tse (ADT model) [7], which pictorially represents information flows and their interactions. The simplicity of these models led to progress on the relay and interference channels. However, the same simplicity prevents them from capturing the interference alignment phenomena observed in Gaussian fading channels. ADT models only represent gains that are powers of two—they are limited in their ability to capture real-valued channel gains.

Real interference alignment exploits channel diversity and channel state information at the transmitter [78]. Channel gains that are not just powers of two are essential to capture it. However, the naive approach with full bit-level multiplication leads to a different problem: the "carries" in the multiplication of the channel gain and the signal break the bit-wise linearity (and hence the simplicity) of the model. To circumvent this issue, the models recently proposed by Maddah-Ali and Niesen [82] ignore these carries to preserve linearity and yet model generic channel gains. This inspires us to call these models "carry-free" models. Essentially, these carry-free models replace a single real channel input by a discrete valued signal and a real channel gain by an LTI system (Fig. 2.2). Sections 2.2 and 2.3 give details of the carry-free interpretation of [82]. Section 2.3.3 deals with MIMO channels.

As a summary, Table 2.1 provides a comparison of the different families of models. While neither ADT nor carry-free models provide direct mappings for individual instances of Gaussian networks, carry-free models are able to capture fractional degrees of freedom and real interference alignment [82].

Armed with this new richer model, Section 2.4 attempts to understand another Gaussian phenomenon: communication over unknown i.i.d. fading channels. A first curiosity about such channels was conjectured in 1969 [94]—even for continuous fading distributions the optimal channel input distribution is discrete. Since then many works have followed up to characterize unknown point-to-point models [108, 1, 63]. Gaussian channel capacities in this case behave as $\log(\log SNR)$, as opposed to the $\log SNR$ scaling for channels with known fading.

Carry-free models prove to be particularly useful in understanding information flows in control theory. Historically, the fields of communication and control have developed independently. Traditional information theory ignores message semantics and focuses on system rate, with complexity, delay and reliability being secondary considerations. On the other hand, traditional control theory has a natural focus on interpreting an observation to perform the right action and cares about the rate of convergence and computable solutions. In the past decade, questions in the area of control over communication networks have forced the two fields to interact. We believe carry-free models help connect these two fields and provide a information-centric perspective. This idea builds on the previous use of ADT models in control problems [47, 84, 93]. We discuss how carry-free models can help understand the information bottlenecks in estimation and control of simple linear systems, in the presence of both explicit rate-limits and implicit uncertainty bottlenecks. Since carry-free models can model unknown fading in communication channels they are particularly useful to model parameter uncertainty in control systems.

The final set of models we introduce remove the randomness from the carry-free models for unknown fading channels. We propose the max-superposition model in Section 2.4 that removes this randomness. The model builds from the observation that only the position of the highest non-zero bit in a signal conveys information when the channel is unknown. The max-model explains how the energy level modulation scheme, which is optimal in the case of real channels [63], can in fact be thought of as simple pulse-position-modulation. A similar fact is true in networks with unknown fading. In the max-superposition model, superimposed signals are combined through the non-linear max-operator, as opposed to a linear sum (hence, the name).

Finally, we note that the max-superposition model of the relay channel offers a surprise: partial-decode-and-forward can offer an unbounded gain over compress-and-forward[1].

2.2 Carry-free operations

Just as in the ADT models [7], the generalized deterministic models of Niesen and Maddah-Ali [82] consider the binary-string representation of real numbers as a starting point. To be able to interpret these models as "carry-free" models we define carry-free arithmetic, i.e. arithmetic without the usual carryovers that result from addition.

Consider such a string $a_n a_{n-1} ... a_1 a_0 a_{-1} ...$. Let $\mathbb{B}$ be the set of all semi-infinite binary strings. We define carry-free addition, $\oplus_c$, and multiplication, $\otimes_c$, over $\mathbb{B}$ in the obvious way.

[1] Most of the results in this chapter are joint work with Se Yong Park [85].

Addition and Multiplication

Let $c_n ... c_1 c_0 c_{-1} ... = a_n ... a_1 a_0 a_{-1} ... \oplus_c b_n ... b_1 b_0 b_{-1} ....$ Then $c_i = a_i + b_i$, where the sum is over $\mathbb{F}_2$. The addition operation involves no carryovers unlike in real addition. We derive the name "carry-free" from this property. Bit interactions at one level do not affect higher level bits.

Let $c_n ... c_1 c_0 c_{-1} ... = a_n ... a_1 a_0 a_{-1} ... \otimes_c b_n ... b_1 b_0 b_{-1} ...$. Then $c_i = \sum_k a_k b_{i-k}$, where the summations are in $\mathbb{F}_2$. Thus, multiplication is basically "convolution" (Fig. 2.2), as noted in [82]. We note that it is commutative and associative.

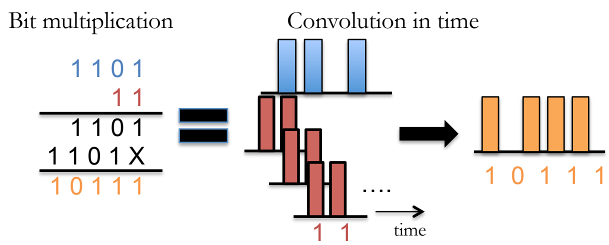


Figure 2.2: Carry-free multiplication as "convolution"

Signal notation

The bit-levels of carry-free numbers are like a fake "time". Signals convolve to interact—just as carry-free strings multiply. The string $a_n ... a_1 a_0 a_{-1} ...$ can also be denoted as signal samples from $-\infty$ to n: $a[n]a[n-1]...a[0]a[-1]....$

Polynomial notation

Under the multiplication and addition defined above, $(\mathbb{B}, \oplus_c, \otimes_c)$ forms a commutative ring. There is a natural mapping from this ring to the polynomial ring. $a_n ... a_1 a_0 a_{-1} ...$ in $\mathbb{B}$ corresponds to the formal power series $a_n z^n + ... + a_1 z + a_0 + a_{-1} z^{-1} ....$ We can check the natural addition and multiplication in the polynomial ring is equivalent to $\oplus_c$ and $\otimes_c$ of $\mathbb{B}$.

Truncation

We will use a truncation operation similar to [7, 2, 14] to capture the effect of noise in a channel. For a formal power series (polynomial) $a(z) = \sum_i a_i z^i$, we define $\lfloor a(z) \rfloor = \sum_{i \geq 0} a_i z^i$.

2.3 Carry-free models

This section walks through the development of the lower triangular or carry-free approximation to standard Gaussian models, similar to that in [7, 82].

2.3.1 The point-to-point channel

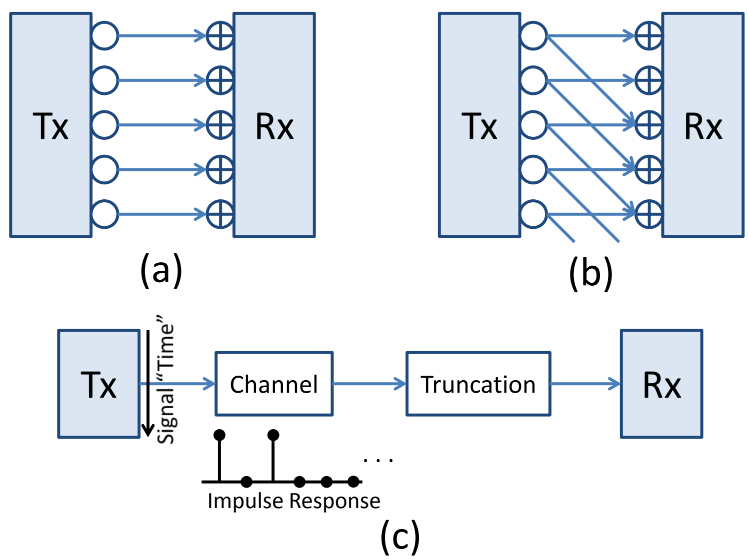


Figure 2.3: (a) ADT model for point-to-point channel with gain $2^5(1+2^{-2})$ (b) The carry-free model for point-to-point channel with gain $2^5(1+2^{-2})$ captures the bit-level structure more finely (c) Signal processing system representation of the carry-free model

Consider a real Gaussian point-to-point channel,

$$y = 2^n \cdot h \cdot x + w. \tag{2.1}$$

h is a fixed realization of a random channel gain uniform in $[0, 1)$. $(2^n)^2$ represents the signal-to-noise ratio (SNR). The noise w is $\mathcal{N}(0, 1)$. For simplicity, we assume signal x has a peak power constraint of 1 as in [7]. We develop a carry-free model for this channel, parallel to the development of the ADT model, and like them consider bounded additive noise instead of Gaussian.

Considering the binary expansions of h and x, we have

$$y = 2^n(h_{-1}2^{-1} + h_{-2}2^{-2} + ...)(x_{-1}2^{-1} + x_{-2}2^{-2} + ...) + (w_{-1}2^{-1} + w_{-2}2^{-2} + ...) \tag{2.2}$$

Replacing the binary expansion with polynomial notation, we get

$$y(z) = z^n(h_{-1}z^{-1} + h_{-2}z^{-2} + ...)(x_{-1}z^{-1} + x_{-2}z^{-2} + ...) + (w_{-1}z^{-1} + w_{-2}z^{-2} + ...), \tag{2.3}$$

We shorten this as $y(z) = z^n h(z)x(z) + w(z)$. For $i \geq 1$, w_{-i} are modeled as Bernoulli$-(\frac{1}{(}2))$ random variables. This captures the idea that any bits of $x(z)$ that are below the noise level are unobservable at the receiver. We can write the channel equation as $y(z) = \lfloor z^n h(z)x(z) \rfloor$. Here, $h(z)$ is the channel polynomial, $\frac{1}{2}\log(\text{SNR}) = n$ and $x(z)$ is the message polynomial. In matrix form this is:

$$\begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{n-3} \\ y_{n-2} \end{bmatrix} = \begin{bmatrix} h_{-1} & h_{-2} & \cdots & h_{-(n-2)} & h_{-(n-1)} \\ 0 & h_{-1} & \cdots & h_{-(n-3)} & h_{-(n-2)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & h_{-1} & h_{-2} \\ 0 & 0 & \cdots & 0 & h_{-1} \end{bmatrix} \begin{bmatrix} x_{-(n-1)} \\ x_{-(n-2)} \\ \vdots \\ x_{-2} \\ x_{-1} \end{bmatrix}, \tag{2.4}$$

or $y = H \cdot x$. Thus, the capacity of the channel is the rank of the $(n-1) \times (n-1)$ matrix H, which is determined by the highest location where h_i is non-zero.

Theorem 2.3.1. *(from [82]) The capacity of the carry-free point-to-point channel with the channel coefficient $h(z) = \sum_i h_i z^i$ and $\frac{1}{2}\log$(SNR) $= n$ is Rank$(H) = n - \min\{i | h_{-i} \neq 0\}$, where H is the matrix as above.*

Remark 2.3.2. The ADT model for a point-to-point channel is a special case of the carry-free model. The ADT model can only capture the behavior of a one-tap impulse response, whereas the carry-free model captures an infinite impulse response with a truncated received signal. Similarly, in the polynomial framework the ADT model allows the channel gain polynomial $h(z)$ to be a monomial, and thus the gain is experienced as a bit shift. The capacity gap between the Gaussian and the ADT model is exactly the same as the gap between the Gaussian and the carry-free model, since in both cases it only depends on the highest bit. Fig. 2.3(a), (b) compare the ADT model to the corresponding carry-free model. Fig. 2.3(c) shows the parallel between bit-levels and "time" in carry-free models.

2.3.2 The multiple-access channel

For completeness we discuss the fractional degrees of freedom in the multiple-access channel as in [82]. We introduce the notion of degrees of freedom of a channel as the achievable rate normalized by $\frac{1}{2}\log(\text{SNR})$. Let $C(h(z), n)$ denote the capacity of the channel with gain vector $h(z)$ and $\frac{1}{2}\log(\text{SNR}) = n$.

Definition 2.3.3. Let $h(z)$ be the channel gain vector. Then the degrees of freedom (d.o.f) of the channel under the carry-free model is defined as

$$d.o.f. = \limsup_{n\to\infty} \frac{C(h(z), n)}{n}. \tag{2.5}$$

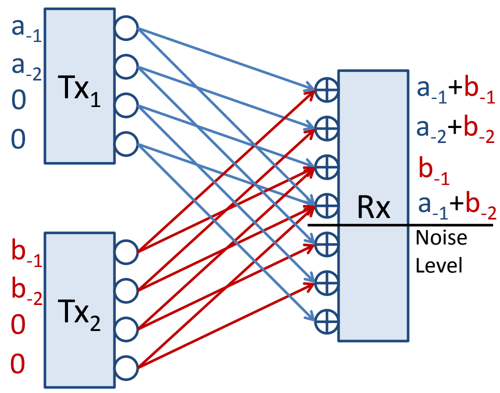


Figure 2.4: MAC with channel gains $h(z) = z^4(1 + z^{-3})$, $h'(z) = z^4(1 + z^{-2})$. $(\frac{1}{2}, \frac{1}{2})$ d.o.f. are achievable using the scheme described.

An interesting question about fractional degrees of freedom was implicitly posed and answered positively by Motahari et al. [78] and Wu at al. [119]: can an encoder achieve $(\frac{1}{2}, \frac{1}{2})$ d.o.f. without time-sharing or CSIT (channel state information at transmitters)? We can answer this question in a carry-free MAC model, as explored in [82] by Niesen and Maddah-Ali [2].

Consider a MAC with no CSIT: $y(z) = \lfloor 2^n h(z)a(z) + 2^n h'(z)b(z) \rfloor$ with $\frac{1}{2}\log(\text{SNR}) = n$. In other words, $h(z)$ and $h'(z)$ are only known to the receiver, and transmitters are not allowed any cooperation or common randomness for time-sharing. Clearly, the cut-set bound implies that the maximum d.o.f. of the channel are bounded by 1. The scheme we use follows [78], and sets the lower order bits to zero as described in Fig. 2.4.

Theorem 2.3.4. *(from [82]) Each transmitter can achieve $\frac{1}{2}$ d.o.f. without timesharing or CSIT for almost all channel gains.*

Proof: For $\epsilon > 0$, let $l = \lfloor \frac{n(1-\epsilon)}{2} \rfloor$. To achieve d.o.f. $(\frac{1}{2}, \frac{1}{2})$ we set the message polynomials as $a(z) = a_{-1}z^{-1} + ... + a_{-l}z^{-l}$ and $b(z) = b_{-1}z^{-1} + ... + b_{-l}z^{-l}$. The relationship between the coefficients received y and the message is then given by eq. (2.6).

Denote eq. (2.6) as $y = H \cdot x$, where x is the stacked vector $\begin{bmatrix} a_{-l} & b_{-l} & \dots & a_{-1}b_{-1} \end{bmatrix}$. To prove the theorem, we will put a random measure on the channel coefficients and prove that for all $\epsilon > 0$ the probability of H being rank-deficient goes to zero. Let $h_{-1}, ..., h'_{-1}, ...$ be

[2]The fundamental issue is that of symmetry — can a symmetric scheme be implemented without pre-coordination? Successive-decoding schemes in Gaussian models are inherently asymmetric. For instance, consider a scheme where one encoder reduces its power to $\sqrt{P}$ while the other transmits at full power P. This requires a pre-coordination on which encoder must reduce power and is equivalent to a time-sharing scheme in cooperation burden. The same must occur in ADT models since the bit shift alignments at the receiver are unknown to the transmitter.

$$\begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{n-l-1} \\ y_{n-l} \\ \vdots \\ y_{n-2} \end{bmatrix} = \begin{bmatrix} h_{l-n} & h'_{l-n} & h_{l-1-n} & h'_{l-1-n} & \cdots & h_{1-n} & h'_{1-n} \\ h_{l+1-n} & h'_{l+1-n} & h_{l-n} & h'_{l-n} & \cdots & h_{2-n} & h'_{2-n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ h_{-1} & h'_{-1} & h_{-2} & h'_{-2} & \cdots & h_{l-1-n} & h'_{l-1-n} \\ 0 & 0 & h_{-1} & h'_{-1} & \cdots & h_{l-2-n} & h'_{l-2-n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & h_{-1} & h'_{-1} \end{bmatrix} \begin{bmatrix} a_{-l} \\ b_{-l} \\ \vdots \\ a_{-1} \\ b_{-1} \end{bmatrix} \tag{2.6}$$

Bernoulli($\frac{1}{2}$) random variables. To guarantee unique decodability, it is enough to show that for all $x \neq 0$, $H \cdot x \neq 0$, since the code is linear.

We do this by partitioning the vectors x into the sets A_1 to A_l based on the first non-zero location. We then use the union bound to bound $\Pr\{Hx = 0\}$. Let $A_1 = \{x | a_{-1} \neq 0 \text{ or } b_{-1} \neq 0\}$. $A_i = \{x | a_{-j} = b_{-j} = 0 \ \forall \ j < i \ , \ a_{-i} \neq 0 \text{ or } b_{-i} \neq 0\}$ for $1 < i \leq l$. The calculations are below.

$$\begin{aligned}
&\Pr(\{Hx = 0 \text{ for some } x \neq 0\}) \\
\leq &\Pr(Hx = 0 \text{ for some } x \in A_1) + \Pr(Hx = 0 \text{ for some } x \in A_2) + \cdots + \Pr(Hx = 0 \text{ for some } x \in A_l) \\
\leq &\max_{x \in A_1} \Pr(Hx = 0) \cdot |A_1| + \max_{x \in A_2} \Pr(Hx = 0) \cdot |A_2| + \cdots + \max_{x \in A_l} \Pr(Hx = 0) \cdot |A_l| \\
= &2^{-(n-1)} \cdot (3 \cdot 2^{2l-2}) + 2^{-(n-2)} \cdot (3 \cdot 2^{2l-4}) + \cdots + 2^{-(n-l)} \cdot 3 \\
\leq &l \cdot 2^{-(n-1)} \cdot (3 \cdot 2^{2l-2}) \\
= &\lfloor \frac{n(1-\epsilon)}{2} \rfloor \cdot 2^{-(n-1)} \cdot (3 \cdot 2^{2\lfloor \frac{n(1-\epsilon)}{2} \rfloor - 2}) \\
&\to 0 \text{ as } n \to \infty
\end{aligned}$$

For eq. 2.7, note that $\max_{x \in A_1} \Pr(Hx = 0)$ is happens in the case where x is 0 in all locations except a_{-1} or b_{-1}. To get $|A_1|$, note that there are 3 configurations of (a_{-1}, b_{-1}) that work, and then there are 2^{2l-2} choices for the other values.

Therefore, the probability that H is full-rank goes to 1 as $n \to \infty$. Hence, we can achieve the d.o.f. $(\frac{1}{2}, \frac{1}{2})$ for almost all channels. ∎

In the above proof, we only use the upper $\frac{n}{2}$ coefficients of the message polynomials. The Gaussian counterpart of this idea would be to increase the quantization bin size to make room for the interference, as in [78]. The proof of this idea in the carry-free model relies on simple matrix invertibility and the union bound, and may offer some insight into the heart of the Khintchine-Groshev theorem used in [78].

Wu et al. [119] suggests the importance of probability distributions that have Rényi dimension $\frac{1}{2}$ to achieve fractional degrees of freedom. One such distribution over numbers involves setting every other bit to zero, while the remaining bit are random, and Cadambe et al. [17] showed that the notion of fixing the values of alternate bits works for the K user interference channel for a restricted set of ADT models. However, since this strategy worked only for the specific set, it was unclear whether the phenomenon was a corner case or was in fact the typical situation in Gaussian channels. We can prove that an alternating bit scheme works for almost all carry-free models in the same way as the above proof. This fact gives an accessible intuition for why Wu et al.'s scheme works.

2.3.3 Modeling point-to-point MIMO

The following 2×2 MIMO example from [7] illustrates a limitation of the ADT model:

$$Y = 2^k \begin{bmatrix} \frac{3}{4} & 1 \\ 1 & 1 \end{bmatrix} X + N \tag{2.7}$$

This channel has 2 d.o.f under the Gaussian model, but only one under ADT models. However, carry-free models represent the channel as $h_{11}(z) = z^{k-1} + z^{k-2},\ h_{12}(z) = z^k,\ h_{21}(z) = z^k, h_{22}(z) = z^k$ and can capture the 2 d.o.f.

Unfortunately, the carry-free model cannot solve this problem generally. There is no simple correspondence between a specific instance of a real 2×2 MIMO model and a specific carry-free model. Consider the channel

$$Y = 2^k \begin{bmatrix} 9 & 3 \\ 3 & 1 \end{bmatrix} X + N, \tag{2.8}$$

with 1 d.o.f. under the real Gaussian model. However, the corresponding carry-free model with $h_{11}(z) = z^3 + 1,\ h_{12}(z) = z + 1,\ h_{21}(z) = z + 1,\ h_{22}(z) = 1$ has 2 d.o.f., since its determinant in the polynomial ring is non-zero. But

$$Y = 2^k \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix} X + N, \tag{2.9}$$

has 2 d.o.f. under the real Gaussian model. Under the carry-free model, the channel coefficients are $h_{11}(z) = z^2 + 1,\ h_{12}(z) = z + 1,\ h_{21}(z) = z + 1,\ h_{22}(z) = 1$, which lead to determinant 0.

This inconsistency between the carry-free model and the Gaussian model seems to arise from the fact that numbers behave differently under the two algebras because of the lack of carryovers. For instance, 5, a prime integer, is composite in the carry-free ring ($101 = 11 \otimes_c 11$). This is the primary difficulty in building a direct map from Gaussian instances to their carry-free counterparts.

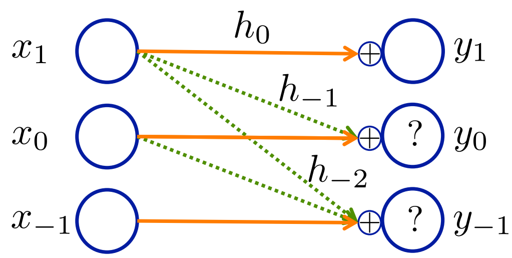


Figure 2.5: This figure shows a carry-free model for an unknown fading channel. The maximal non-zero bit of the input contaminates output at all lower-levels. $y_1 = h_0 x_1$ remains correlated to $x(z)$. However, $y_0 = h_0 x_0 + h_{-1} x_1$, is now a Bernoulli$-\frac{1}{2}$ random variable independent of $x(z)$ if $x_1 = 1$. The same is true for all lower levels of $y(z)$.

2.4 Channels with unknown i.i.d. fading

2.4.1 Point-to-point carry-free model

Consider the carry-free model for a channel with unknown i.i.d. fading as in Fig. 2.5. The basic setup builds from earlier sections[3]: $y(z) = \lfloor h(z)x(z) \rfloor$ where $h(z) = h_0 + h_{-1}z^{-1} + \cdots$ and $x(z) = x_n z^n + x_{n-1}z^{n-1} + \cdots + x_0$. Here, h_i are i.i.d. Bernoulli$-\frac{1}{2}$ random variables which are unknown to both the transmitter and the receiver, and $x_i \in \{0, 1\}$.

The first thing to note is that this carry-free channel is not deterministic—the randomness from the fading introduces randomness in the input-output relationship. Next, observe that once the maximum non-zero bit-level in the output is fixed, the randomness introduced by the channel forces all bits below the maximum level to be Bernoulli-$\frac{1}{2}$ random variables which are independent from the input. Therefore, we focus our attention on the highest 1 received in $y(z)$. Define,

$$y_q = \begin{cases} -1 & \text{if } y(z) = 0 \\ \max\{i | y_i = 1\} & \text{if } y(z) \neq 0 \end{cases} \tag{2.10}$$

Similarly we can also define the maximal bit of $x(z)$ as,

$$x_q = \begin{cases} -1 & \text{if } x(z) = 0 \\ \max\{i | x_i = 1\} & \text{if } x(z) \neq 0 \end{cases} \tag{2.11}$$

Given that $y(z) = \lfloor h(z)x(z) \rfloor$ and the degree of $h(z)$ is 0, $x_q \geq y_q$. The key observation is that x_q and y_q are sufficient statistics of $x(z), y(z)$ for $y(z), x(z)$ respectively, and this is captured in the following lemma.

[3]For ease of exposition we adopt slightly different model parameters and incorporate the channel power gain z^{n+1} (not z^n) into $x(z)$ and $h(z)$.

Theorem 2.4.1. *$P(y_q = x_q - i) = \left(\frac{1}{2}\right)^{i+1}$ for $0 \le i \le x_q$ and $P(y_q = -1) = \left(\frac{1}{2}\right)^{x_q+1}$. Moreover, $x(z)$—x_q—y_q—$y(z)$ form a Markov chain, and $I(y(z); x(z)) = I(y_q; x(z)) = I(y_q; x_q)$.*

Proof: $y_{x_q} = h_0 \cdot x_{x_q}$. Hence, $P(y_q = x_q) = P(h_0 = 1) = \frac{1}{2}$. Similarly,

$$y_{x_q-i} = \sum_{j=0}^{x_q} h_{j-i} \cdot x_{x_q-j} \tag{2.12}$$

$P(y_q = x_q - i) = P(h_0 = 0, h_{-1} = 0, \cdots h_{-i+1} = 0) = \left(\frac{1}{2}\right)^n$.

$$\begin{aligned}
I(y(z); x(z)) &= H(y(z)) - H(y(z)|x(z)) && (2.13)\\
&= H(y_{y_q}, y_{y_q-1} \dots y_0) - H(y_{y_q}, y_{y_q-1} \dots y_0 | x_{x_q}, x_{x_q-1} \dots x_0) && (2.14)\\
&= H(y_{y_q}) + H(y_{y_q-1}, y_{y_q-2} \dots y_0 | y_{y_q}) - H(y_{y_q} | x_{x_q}, x_{x_q-1x+0} \dots) \\
&\quad - H(y_{y_q-1}, y_{y_q-2} \dots | y_{y_q}, x_{x_q}, x_{x_q-1} \dots) && (2.15)\\
&= I(y_{y_q}; x_{x_q}, x_{x_q-1} \dots x_0) + H(y_{y_q-1}, y_{y_q-2} \dots y_0 | y_{y_q}) \\
&\quad - H(y_{y_q-1}, y_{y_q-2} \dots y_0 | y_{y_q}, x_{x_q}, x_{x_q-1} \dots x_0) && (2.16)\\
&= I(y_{y_q}; x(z)) - I(y_{y_q-1}, y_{y_q-2} \dots y_0; x_{x_q}, x_{x_q-1} \dots x_0 | y_{y_q}) && (2.17)\\
&= I(y_{y_q}; x(z)) && (2.18)
\end{aligned}$$

We must show that

$$I(y_{y_q-1}, y_{y_q-2} \dots y_0; x_{x_q}, x_{x_q-1} \dots x_0 | y_{y_q}) = 0 \tag{2.19}$$

to get eq. (2.18).

$$I(y_{y_q-1}, y_{y_q-2} \dots y_0; x_{x_q}, x_{x_q-1} \dots x_0 | y_{y_q}) = \sum_{i=0}^{y_q-1} I(y_i; x_{x_q}, x_{x_q-1} \dots x_0 | y_{y_q}, y_{y_q-1} \dots y_{i+1}) \tag{2.20}$$

Now, we note, since $x_q \ge y_q$ we can write:

$$\begin{aligned}
y_{y_q-i} &= h_{y_q-i-x_q} \cdot x_{x_q} + \sum_{j=0}^{x_q-1} h_{y_q-i-x_q+j} x_{x_q-j} && (2.21)\\
&= h_{y_q-i-x_q} \cdot 1 + \sum_{j=0}^{x_q-1} h_{y_q-i-x_q+j} x_{x_q-j} && (2.22)
\end{aligned}$$

So y_{y_q-i} is independent of all $x_i, i < x_q$ due to the Bernoulli$-\frac{1}{2}$ additive random variable $h_{y_q-i-x_q}$. $x_q \geq y_q$ guarantees that $h_{y_q-i-x_q}$ is included in the sum. So conditioned on y_{y_q} all lower bits are independent of all input bits and we have:

$$I(y_{y_{y_q-1}}; x_{x_q}, x_{x_q-1} \dots x_0 | y_{y_q}, y_{y_q-1} \dots y_{y_q-i+1}) = 0 \tag{2.23}$$

which gives the required result eq. (2.19).

$$\begin{aligned} I(y_{y_q}; x(z)) &= H(y_q) - H(y_{y_q}|x_{x_q}, x_{x_q-1} \dots x_0) &(2.24)\\ &= H(y_q) - H(y_{y_q}|x_{x_q}) &(2.25)\\ &= I(y_{y_q}; x_{x_q}) &(2.26)\\ &= I(y_q; x_q) &(2.27)\end{aligned}$$

For eq. (2.25), $H(y_{y_q}|x_{x_q}, x_{x_q-1} \dots x_0) = H(y_{y_q}|x_{x_q})$ since $y_{y_q} = h_{y_q-x_q} \cdot x_{x_q}$, where $h_{y_q-x_q} = max_i\{h_i = 1\}$ is the highest 1 in $c(z)$. So bits lower than x_q are independent of y_{y_q}, since all $h_i, i > y_q - x_q$ are 0. For eq. (2.26), the only randomness in $x_{x_q} = 1$ and $y_{y_q} = 1$ comes from their indices x_q, y_q.

■

We use the lemma to calculate the capacity of this channel, $C_{cf} = \max_{p(x(z))} I(x(z); y(z))$.

Theorem 2.4.2. *The capacity of the carry-free point-to-point channel with unknown fading is* $\log(n+2) - 3 \leq C_{cf} \leq \log(n+2)$.

Proof: The achievable scheme involves using monomials z^{x_q} as the channel input $x(z)$ and parallels pulse-position modulation. The converse follows from Lemma 2.4.1. ■

Remark 2.4.3. This behavior does not manifest in an ADT model. Consider a point-to-point ADT model where the degree of the gain h is drawn uniformly from 0 to n and is unknown to the transmitter and the receiver. Let x_q be the position of the highest 1 on the transmitted signal, and let h_q be the realization of the degree of the gain. Then the receiver can only see $x_q + h_q$. Since both these quantities are unknown, it is not possible to decrease the entropy of x_q below $\log n$ based on the received signal, and this model predicts that no information could be transmitted across the channel.

No, what about the case where the channel gain h has degree larger than 1 with more than 1 deterministic bits (e.g. Fig. 2.6)? This is discussed in the next corollary.

Corollary 2.4.4. *Let*

$$\begin{aligned} h(z) &= h_g \cdot z^g + h_{g-1}[n] \cdot z^{g-1} + \cdots &(2.28)\\ &= 1 \cdot z^{g_{det}} + 0 \cdot z^{g_{det}-1} + 0 \cdot z^{g_{det}-2} + \cdots \\ &\quad + h_{g_{rand}} \cdot z^{g_{rand}} + h_{g_{rand}-1} \cdot z^{g_{rand}-1} + \cdots &(2.29)\end{aligned}$$

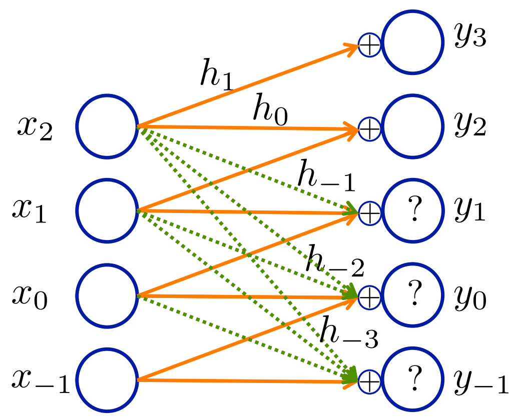


Figure 2.6: Here both h_1 and h_0 are known. $h_1 = 1$ and $h_0 = 1$ say. Then both x_2 and x_1 can be decoded from $y(z)$.

where

$$g_{rand} = \max\{i | c_i \sim \mathit{Bernoulli}(\frac{1}{2}), g_c \geq i \geq -\infty\}. \tag{2.30}$$

and

$$g_{det} = g. \tag{2.31}$$

If $y(z) = \lfloor h(z)x(z) \rfloor$, *then* $I(y(z); x(z)) \leq g_{det} - g_{rand} + \log n$ *where* n *is the maximum possible degree of* $x(z)$.

Proof: The proof follows exactly the same as that of the previous lemma.

$$I(x(z); y(z)) = I(x(z); y_{x_q+g_{det}}, y_{x_q+g_{det}-1}, \cdots, y_{x_q+g_{rand}}, \cdots y_0) \tag{2.32}$$

$$y_{x_q+g_{rand}-i} = \sum_{j=0}^{g_{det}-g_{rand}} h_{g_{rand}+j-i} \cdot x_{x_q-j} \tag{2.33}$$

$$= h_{g_{rand}-i} x_{x_q} + \sum_{j=1}^{g_{det}-g_{rand}} h_{g_{rand}+j-i} \cdot x_{x_q-j} \tag{2.34}$$

Since $h_{g_{rand}-i} x_{x_q}$ is a Bernoulli$-(\frac{1}{2})$ random variable, $y_{x_q+g_{rand}-i}$ is independent of x_i for all i. So we have

$$I(x(z); y(z)) = I(x(z); y_{x_q+g_{det}}, y_{x_q+g_{det}-1}, \cdots, y_{x_q+g_{rand}+1}) \tag{2.35}$$

$$\leq H(y_{x_q+g_{det}}, y_{x_q+g_{det}-1}, \cdots, y_{x_q+g_{rand}+1}) \tag{2.36}$$

$$= g_{det} - g_{rand} + \log n \tag{2.37}$$

$g_{det} - g_{rand}$ bits of $x(z)$ can be decoded from $y_{x_q+g_{det}}$ to $y_{x_q+g_{rand}+1}$. The position x_q is a deterministic function of $y_{x_q+g_{det}}$.

■

2.4.2 Point-to-point max-superposition

The observations above suggest that the essence of communicating over unknown fading channels is captured by the location of the highest non-zero bits of the input and output. Hence, we propose the max-model, which depends only on the maximal bit of the signal. Let $x \in \{0, 1, 2 \cdots n\}$ represent the position of the maximal bit in the input signal. Then, y, the position of the maximal bit in the output signal is represented as simply as

$$y = x. \tag{2.38}$$

Theorem 2.4.5. *The capacity of this channel is* $C_{max} = \log(n + 1)$.

Proof: Since the channel is noiseless and the input can take $n + 1$ values, clearly $C_{max} = \log(n + 1)$. ■

2.4.3 Point-to-point real model

The carry-free models give an intuitive explanation of the $\log \log SNR$ result for communication over unknown fading channels. For a channel with uniform distribution: $Y = H \cdot X + W$, unknown i.i.d. fading $H \sim U[0, 1]$, input power constraint $\mathbb{E}[X^2] \leq P$ and noise $W \sim U[0, 1]$.

We know from [63] that the optimal achievable scheme for this channel is to modulate the energy-level, and there are $n = \log P$ of those available (i.e. the number of bit-levels in the carry-free model). The different signal values (logarithmically spaced) that are used as inputs can be thought of as pulse-positions for standard pulse-position-modulation. The observation that $\lfloor \log |Y| \rfloor^+$ is an approximate sufficient statistic for Y, clearly illustrates the insight and applicability of the max-model to this problem.

2.4.4 The carry-free MAC with unknown fading

Next we move to the carry-free unknown MAC. User A has message $a(z)$ and User B has message $b(z)$. $y(z) = \lfloor h(z)a(z) + h'(z)b(z) \rfloor$ where $h(z) = h_0 + h_{-1}z^{-1} + \cdots$ and $h'(z) = h'_0 + h'_{-1}z^{-1} + \cdots$ are unknown at the transmitter and receiver. Let $a(z) = a_{n_1}z^{n_1} + \cdots + a_0$ and $b(z) = b_{n_2}z^{n_2} + \cdots + b_0$, and y_q, a_q, and b_q are defined as in the previous section.

Just as in the point-to-point case, the randomness introduced by the fading restricts the information communicated across the channel to the maximal power level over the two input signal, and y_q depends only on $\max\{a_q, b_q\}$. This intuition is captured in Lemma 2.4.6.

Lemma 2.4.6. *$P(y_q = \max\{a_q, b_q\} - i) = \left(\frac{1}{2}\right)^{i+1}$ for $0 \leq i \leq \max\{a_q, b_q\}$ and $P(y_q = -1) = \left(\frac{1}{2}\right)^{\max\{a_q,b_q\}+1}$. Further, $a(z), b(z)$—a_q, b_q—$\max\{a_q, b_q\}$—y_q—$y(z)$ form a Markov chain. Hence, $I(y(z); a(z), b(z)) = I(y_q; a(z), b(z)) = I(y_q; a_q, b_q)$.*

Proof: This proof follows exactly as the proof of Lemma 2.4.1. ■

Just as in the point-to-point case, Lemma 2.4.6 suggests a deterministic model, described in the next section.

2.4.5 The max-superposition MAC

Finally, we come to the example which gives the max-superposition model its name. Let $a \in \{0, 1, 2 \cdots n_1\}$ and $b \in \{0, 1, 2 \cdots n_2\}$ represent the highest non-zero bits of the two input signals in a MAC that carry the messages of user A and user B respectively. Then the output, y, of the channel depends only on the maximum of the two signals:

$$y = \max\{a, b\}. \tag{2.39}$$

The very nature of this model suggests that it is not possible for both encoders to simultaneously communicate information. We believe this also captures the behavior in the AWGN MAC with unknown gains. Theorem 2.4.7 below proves that the timesharing rate region is optimal to within one bit.

Theorem 2.4.7. *The achievable rate region for the max-superposition MAC consists of all positive rate pairs (R_1, R_2) such that*

$$\log(n_2 + 1)R_1 + \log(n_1 + 1)R_2 \leq \log(n_1 + 1)\log(n_2 + 1) \tag{2.40}$$

Further, this region is tight to within 1 bit for both R_1 and R_2 — namely all achievable pairs must satisfy:

$$\log(n_2 + 1)(R_1 - 1) + \log(n_1 + 1)(R_2 - 1) \leq \log(n_1 + 1)\log(n_2 + 1)$$

The achievable rate region is a simple timesharing scheme. For the converse we use the converse for the DMC MAC and evaluate bounds using the two lemmas below. The proofs are given in the Appendix 2.5.

Lemma 2.4.8. *For the max-superposition MAC with channel inputs $X_1, X_2 \in \{0, 1, 2 \cdots n\}$ that are independent, and with $Y = \max(X_1, X_2)$,*

$$I(X_1; Y|X_2) + I(X_2; Y)|X_1) \leq \log(n + 1) + 1 \tag{2.41}$$

Lemma 2.4.9. *Let $X_1 \in \{0, 1 \cdots n_1\}$ and $X_2 \in \{0, 1 \cdots n_2\}$ be independent, and $n_1 > n_2$. Let $Y = \max(X_1, X_2)$. Then there exists $0 \leq p \leq 1$ and $Z_1 \in \{0, 1 \cdots n_2\}$ independent of X_2 such that*

$$I(X_1; Y|X_2) \leq (1-p) \cdot \log(n_1 + 1) + p \cdot I(X_2; \max(Z_1, X_2)|Z_1)) + 1 \tag{2.42}$$

$$I(X_2; Y|X_1) \leq p \cdot I(Z_1; \max(Z_1, X_2)|X_2)) + 1 \tag{2.43}$$

$$I(X_2; \max(Z_1, X_2)|Z_1)) + I(Z_1; \max(Z_1, X_2)|X_2)) \leq \log(n_2 + 1) + 1 \tag{2.44}$$

2.4.6 Relay network

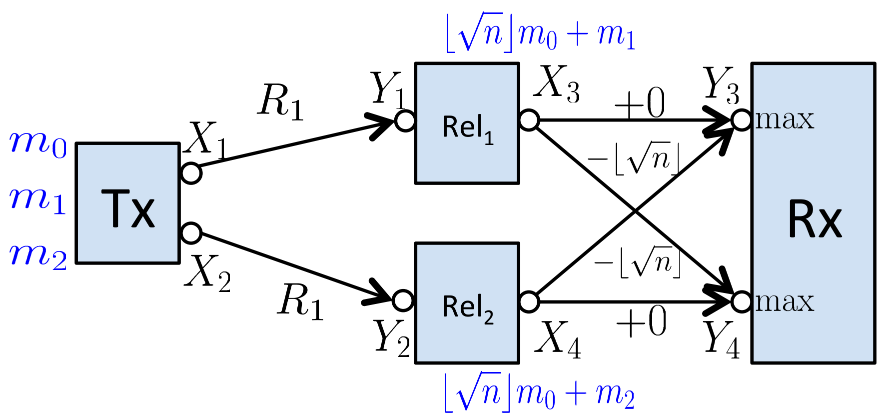


Figure 2.7: A max-superposition relay network with an approximately optimal partial-decode-and-forward scheme. The cross-links between X_3, X_4 and Y_3, Y_4 are weaker than the direct links and lead to downward bit-level shifts by $\lfloor\sqrt{n}\rfloor$. $X_1 = \lfloor\sqrt{n}\rfloor m_0 + m_1$ and $X_2 = \lfloor\sqrt{n}\rfloor m_0 + m_2$ are transmitted by the two relays.

Having understood point-to-point links as well as multiple-access channels with unknown channel gains, we move to networks with unknown fading. In particular, we consider a relay-network here.

Successive decoding is not possible in a MAC with unknown fading — the best scheme is essentially time-sharing. This suggests that there might not be any point to try and understand coordination or beam-forming style gains by relays in more complex networks. Compress-and-forward would seem like the right approximately optimal scheme.

This section gives an example of a relay network with unknown fading where surprisingly, partial-decode-and-forward can infinitely outperform compress-and-forward style schemes. There is a significant coordination gain (traditionally known as beam-forming gain) to be exploited. This observation is quite interesting when we remind ourselves that the beamforming gain is usually bounded by a constant in known-fading Gaussian networks with linear superposition, and compress-and-forward is the optimal in linear ADT relay networks [7]. [4]

In the previous section, we saw that the max operation essentially captures superposition in i.i.d. unknown fading networks. We focus on a specific max-superposition relay network shown in Fig. 2.9. Channels connecting X_1 to Y_1 and X_2 to Y_2 are orthogonal noiseless links with rate R each. The MIMO channel connecting X_3, X_4 to Y_3, Y_4 is a max-superposition deterministic model. We consider $X_3 \in \{0, 1, \cdots, n\}$, $X_4 \in \{0, 1, \cdots, n\}$ and

$$Y_3 = \max(X_3, (X_4 - \lfloor\sqrt{n}\rfloor)^+) \tag{2.45}$$

$$Y_4 = \max((X_3 - \lfloor\sqrt{n}\rfloor)^+, X_4) \tag{2.46}$$

The shifted max-operation captures the difference between the direct-link and cross-link gains in the original Gaussian MIMO channel. The cross links are weaker than the direct links, but not by much in a proper scaling.

First, observe that when $X_3 - X_4 \geq \sqrt{n}$, $Y_3 = X_3$ and $Y_4 = (X_3 - \lfloor\sqrt{n}\rfloor)^+$. Thus, both the received signals only depend on X_3, and X_4 is essentially useless. Moreover, when X_3 and X_4 are independent and their ranges are large, with high probability the difference between X_3 and X_4 is greater than $\sqrt{n}$ and both Y_3 and Y_4 only depend on either X_3 or X_4. Therefore, in compress-and-forward (where X_3 and X_4 are chosen independently), the throughput is essentially bounded by $\log n$ which is the maximum entropy of a single input.

However, decode-and-forward can exploit the coordination gain between X_3 and X_4. We can keep the difference within $\sqrt{n}$ and let Y_3, Y_4 depend on both X_3, X_4. The key observation is that when the quotients of X_3, X_4 divided by $\lfloor\sqrt{n}\rfloor$ are the same, X_3, X_4 have to be within $\sqrt{n}$. To keep the quotients the same, we transfer a common message m_0 to both relays as the common quotient.

An approximately optimal partial-decode-and-forward scheme is described in Fig. 2.9. The transmitter divides the message into three parts $m_0, m_1, m_2 \in \{0, 1, \cdots, \lfloor\sqrt{n}\rfloor - 1\}$. The common message m_0 is sent to both relays and used as a quotient. m_0 essentially is like synchronization and helps to set the appropriate levels for both transmitters. This coordination is key for the scheme. The private messages m_1 and m_2 are transmitted to Rel_1 and Rel_2 respectively and used as remainders. In other words, X_3 and X_4 are chosen as $\lfloor\sqrt{n}\rfloor m_0 + m_1$ and $\lfloor\sqrt{n}\rfloor m_0 + m_2$ respectively. Since the signals are within $\sqrt{n}$, Y_3 and Y_4 also have to equal X_3 and X_4 (eqs. (2.45), (2.46)). Moreover, since the ranges of m_0, m_1, m_2 are $\lfloor\sqrt{n}\rfloor$, the achievable rate of this partial-decode-and-forward can increase to $\frac{3}{2}\log n$.

We summarize our understanding in Thm. 2.4.10 and Fig. 2.8. The proof is included in the Appendix 2.5.

[4] It was already known that partial-decode-and-forward is also optimal in semi-deterministic single-relay channels [61].

Theorem 2.4.10. *Let C_{rel} be the capacity of the relay channel given as above. The proposed partial-decode-and-forward scheme gives the following lower bound.*

$$C_{rel} \geq \min\{2R, R + \log(\lfloor\sqrt{n}\rfloor), 3\log(\lfloor\sqrt{n}\rfloor)\} \tag{2.47}$$

The cut-set bound gives the upperbound

$$C_{rel} \leq \min\{2R, R + \frac{1}{2}\log(n+1) + 2\log 3, \frac{3}{2}\log(n+1) + \frac{3}{2}\log 2\} \tag{2.48}$$

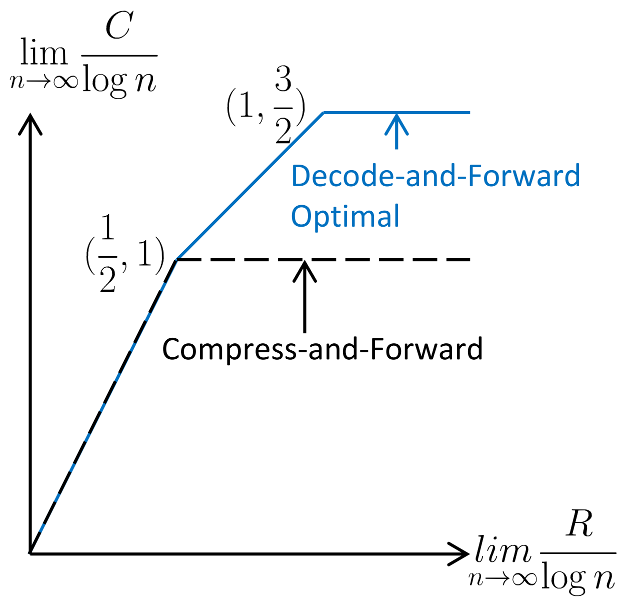


Figure 2.8: Generalized d.o.f. plot for the relay network in Fig. 2.9. $\frac{3}{2}$ d.o.f. capacity is only achievable by partial-decode-and-forward, not by compress-and-forward. Therefore, the gap between the two goes to infinity as n grows.

Hence, the achievable scheme is approximately optimal within a constant gap. Since compress-and-forward cannot exploit a "beam-forming gain", it can perform infinitely worse than the partial-decode-and-forward scheme.

2.5 Appendix

2.5.1 Proof of Lemma 2.4.8

Lemma 2.4.8. *For the max-superposition MAC with channel inputs $X_1, X_2 \in \{0, 1, 2 \cdots n\}$ that are independent, and with $Y = \max(X_1, X_2)$,*

$$I(X_1; Y|X_2) + I(X_2; Y)|X_1) \leq \log(n+1) + 1 \tag{2.41}$$

Proof:

$$
\begin{aligned}
& I(X_1;Y|X_2) + I(X_2;Y|X_1) && (2.49)\\
& = I(X_1;Y|X_2) + I(X_2;Y) - I(X_2;Y) + I(X_2;Y|X_1) && (2.50)\\
& = I(X_1,X_2;Y) - I(X_2;Y) + I(X_2;Y|X_1) && (2.51)\\
& = I(X_1,X_2;Y) - H(X_2) + H(X_2|Y) + H(X_2|X_1) - H(X_2|Y,X_1) && (2.52)\\
& = I(X_1,X_2;Y) - H(X_2) + H(X_2|Y) + H(X_2) - H(X_2|Y,X_1) && (2.53)\\
& \text{(since } X_1 \text{ and } X_2 \text{are independent)} && \\
& = I(X_1,X_2;Y) + H(X_2|Y) - H(X_2|Y,X_1) && (2.54)\\
& = I(X_1,X_2;Y) + I(X_1;X_2|Y) && (2.55)\\
& \leq \log(n+1) + 1 && (2.56)
\end{aligned}
$$

The intuition for the bound on $I(X_1;X_2|Y)$ is simple. Conditioned on the value of the max of two numbers, the mutual information between them is just one bit: who is bigger.

$$
\begin{aligned}
I(X_1;X_2|Y) &= I(X_1;X_2|\max(X_1,X_2)) && (2.57)\\
&= H(X_1|\max(X_1,X_2)) - H(X_1|\max(X_1,X_2),X_2) && (2.58)\\
&\leq H(X_1|\max(X_1,X_2)) - H(X_1|\max(X_1,X_2),X_2,\mathbf{1}_{(X_1\geq X_2)}) && (2.59)\\
&\text{(conditioning reduces entropy)} && \\
&= I(X_2;X_1,\mathbf{1}_{(X_1\geq X_2)}|\max(X_1,X_2)) && (2.60)\\
&= I(\mathbf{1}_{(X_1\geq X_2)};X_2|\max(X_1,X_2)) + I(X_1;X_2|\mathbf{1}_{(X_1\geq X_2)},\max(X_1,X_2)) && (2.61)\\
&\leq 1 + 0 && (2.62)\\
&\text{(The conditioning of the second term reveals either } X_1 \text{ or } X_2\text{)} &&
\end{aligned}
$$

■

2.5.2 Proof of Lemma 2.4.9

Lemma 2.4.9. *Let $X_1 \in \{0,1\cdots n_1\}$ and $X_2 \in \{0,1\cdots n_2\}$ be independent, and $n_1 > n_2$. Let $Y = \max(X_1,X_2)$. Then there exists $0 \leq p \leq 1$ and $Z_1 \in \{0,1\cdots n_2\}$ independent of X_2 such that*

$$
\begin{aligned}
& I(X_1;Y|X_2) \leq (1-p)\cdot\log(n_1+1) + p\cdot I(X_2;\max(Z_1,X_2)|Z_1)) + 1 && (2.42)\\
& I(X_2;Y|X_1) \leq p\cdot I(Z_1;\max(Z_1,X_2)|X_2)) + 1 && (2.43)\\
& I(X_2;\max(Z_1,X_2)|Z_1)) + I(Z_1;\max(Z_1,X_2)|X_2)) \leq \log(n_2+1) + 1 && (2.44)
\end{aligned}
$$

Proof:
Let $p = P(X_1 \leq n_2)$.

Now consider,

$$
\begin{aligned}
& I(X_1; Y|X_2) && (2.63)\\
& = H(Y|X_2) - H(Y|X_2, X_1) && (2.64)\\
& = H(Y|X_2) && (2.65)\\
& = H(Y, \mathbf{1}(X_1 > n_2)|X_2) - H(\mathbf{1}(X_1 > n_2)|X_2, Y) && (2.66)\\
& \leq H(Y, \mathbf{1}(X_1 > n_2)|X_2) && (2.67)\\
& = H(\mathbf{1}(X_1 > n_2)|X_2) + H(Y|X_2, \mathbf{1}(X_1 > n_2)) && (2.68)\\
& \leq 1 + H(Y|X_2, \mathbf{1}(X_1 > n_2) = 1)P(X_1 > n_2) \\
& \quad + H(Y|X_2, \mathbf{1}(X_1 > n_2) = 0)P(X_1 \leq n_2) && (2.69)\\
& \leq 1 + (1-p) \cdot \log(n_1 + 1) + p \cdot H(Y|X_2, \mathbf{1}(X_1 > n_2) = 0) && (2.70)\\
& = 1 + (1-p) \cdot \log(n_1 + 1) + p \cdot H(\max(X_2, Z_1)|X_2) && (2.71)\\
& \quad \text{(since } X_1 \leq n_2 \text{ we can replace it by } Z_1 \text{ which is uniform in } \{0, 1, \ldots n_2\}) \\
& = 1 + (1-p) \cdot \log(n_1 + 1) + p\,(H(\max(X_2, Z_1)|X_2) - H(\max(X_2, Z_1)|X_2, Z_1)) && (2.72)\\
& \quad \text{(because } H(\max(X_2, Z_1)|X_2, Z_1) = 0) \\
& = 1 + (1-p) \cdot \log(n_1 + 1) + p \cdot I(\max(X_2, Z_1), Z_1|X_2) && (2.73)
\end{aligned}
$$

$$
\begin{aligned}
I(X_2; Y|X_1) &= H(Y|X_1) - H(Y|X_1, X_2) && (2.74)\\
&= H(Y|X_1) && (2.75)\\
&\leq H(\max(X_1, X_2)|X_1, \mathbf{1}(X_1 > n_2)) && (2.76)\\
&\quad \text{(conditioning reduces entropy)} && (2.77)\\
&= H(\max(X_1, X_2)|X_1, \mathbf{1}(X_1 > n_2) = 1) \cdot P(X_1 > n_2) && (2.78)\\
&\quad + H(\max(X_1, X_2)|X_1, \mathbf{1}(X_1 > n_2) = 0) \cdot P(X_1 \leq n_2) && (2.79)\\
&= p \cdot H(\max(X_1, X_2)|X_1, \mathbf{1}(X_1 > n_2) = 0) && (2.80)\\
&= p \cdot (H(\max(Z_1, X_2)|Z_1) - H(\max(Z_1, X_2)|Z_1, X_2)) && (2.81)\\
&\quad \text{(since } X_1 \leq n_2 \text{ we can replace it by } Z_1 \text{ which is uniform in } \{0, 1, \ldots n_2\}) \\
&= I(\max(Z_1, X_2); X_2|Z_1) \cdot P(X_1 \leq n_2) && (2.82)
\end{aligned}
$$

Finally, we can use Lemma 2.4.8 for Z_1 and $X_2 \in \{0, 1, \ldots n_1\}$ implies that

$$
I(X_2; \max(Z_1, X_2)|Z_1)) + I(Z_1; \max(Z_1, X_2)|X_2)) \leq \log(n_2 + 1) + 1 \qquad (2.83)
$$

■

2.5.3 Proof of Thm. 2.4.7

Theorem 2.4.7. *The achievable rate region for the max-superposition MAC consists of all positive rate pairs* (R_1, R_2) *such that*

$$\log(n_2+1)R_1 + \log(n_1+1)R_2 \leq \log(n_1+1)\log(n_2+1) \tag{2.40}$$

Further, this region is tight to within 1 *bit for both* R_1 *and* R_2 *— namely all achievable pairs must satisfy:*

$$\log(n_2+1)(R_1-1) + \log(n_1+1)(R_2-1) \leq \log(n_1+1)\log(n_2+1)$$

Proof: Consider a timesharing scheme where time t_1 slots are used by user A and t_2 slots are used by user B, with $t_1 + t_2 = 1$. Then user A achieves rate $R_1 = t_1 \log(n_1+1)$, and user B achieves rate $R_2 = t_2 \log(n_2+1)$. Then we have that $\frac{R_1}{\log(n_1+1)} + \frac{R_2}{\log(n_2+1)} = 1$, which gives the achievable region.

For the converse, we use the converse for a DMC MAC [25]. With channel input X_1, X_2 and output Y we have:

$$R_1 \leq I(X_1; Y|X_2) \tag{2.84}$$
$$R_2 \leq I(X_2; Y|X_1) \tag{2.85}$$
$$R_1 + R_2 \leq I(X_1, X_2; Y) \tag{2.86}$$

We know from Lemma 2.4.9 that there exists $0 \leq p \leq 1$ and $Z_1 \in \{0, 1 \cdots n_2\}$ independent of X_2 such that

$$I(X_1; Y|X_2) \leq (1-p) \cdot \log(n_1+1) + p \cdot I(X_2; \max(Z_1, X_2)|Z_1) + 1 \tag{2.87}$$
$$I(X_2; Y|X_1) \leq p \cdot I(Z_1; \max(Z_1, X_2)|X_2) \tag{2.88}$$
$$I(X_2; \max(Z_1, X_2)|Z_1) + I(Z_1; \max(Z_1, X_2)|X_2) \leq \log(n_2+1) + 1 \tag{2.89}$$

Let $I(Z_1; \max(Z_1, X_2)|X_2) = C_1$ and $I(X_2; \max(Z_1, X_2)|Z_1) = C_2$. We can rewrite the inequalities above as:

$$I(X_1; Y|X_2) \leq (1-p) \cdot \log(n_1+1) + p \cdot C_2 + 1 \tag{2.90}$$
$$I(X_2; Y|X_1) \leq p \cdot C_1 \tag{2.91}$$
$$C_2 + C_1 \leq \log(n_2+1) + 1 \tag{2.92}$$

Now consider $\log(n_2+1)R_1 + \log(n_1+1)R_2$, and apply eq. (2.90) and eq. (2.91)

$$
\begin{aligned}
&\log(n_2+1)R_1 + \log(n_1+1)R_2 && (2.93)\\
&\leq \log(n_2+1)\left[(1-p)\cdot\log(n_1+1) + p\cdot C_2 + 1\right] + \log(n_1+1)\left[p\cdot C_1\right] && (2.94)\\
&= (1-p)\log(n_1+1)\log(n_2+1) + p[C_2\log(n_2+1) + C_1\log(n_1+1)] + \log(n_2+1) && (2.95)\\
&= (1-p)\log(n_1+1)\log(n_2+1) + p[C_2\log(n_1+1) + C_1\log(n_1+1)] + \log(n_2+1) && (2.96)\\
&\quad (\text{ since } n_2 < n_1)\\
&= (1-p)\log(n_1+1)\log(n_2+1) + p\log(n_1+1)\left[C_2 + C_1\right] + \log(n_2+1) && (2.97)\\
&\leq (1-p)\log(n_1+1)\log(n_2+1) + p\log(n_1+1)\left[\log(n_2+1) + 1\right] + \log(n_2+1) && (2.98)\\
&\quad (\text{by eq. } (2.92))\\
&= \log(n_2+1)\log(n_1+1) + \log(n_2+1) + \log(n_1+1) && (2.99)
\end{aligned}
$$

Rearranging this, we get:

$$\log(n_2+1)(R_1-1) + \log(n_1+1)(R_2-1) \leq \log(n_1+1)\log(n_2+1)$$

■

2.5.4 Proof of Thm. 2.4.10

Before we go into the proof of Thm. 2.4.10, we state and prove a few lemmas.
The first lemma establishes an upperbound on the information that can flow between the relays and the receiver.

Lemma 2.5.1. *Let* $X_1, X_2 \in \{0, 1, \cdots, n\}$ *and* $Y_1 = \max\{X_1, (X_2 - \lfloor\sqrt{n}\rfloor)^+\}$, $Y_2 = \max\{(X_1 - \lfloor\sqrt{n}\rfloor)^+, X_2\}$. *Then,*

$$\max_{p(X_1,X_2)} \min\{I(X_1;Y_1,Y_2|X_2), I(X_2;Y_1,Y_2|X_1)\} \leq \frac{1}{2}\log(n+1) + 2\log 3 \qquad (2.100)$$

Proof: We divide X_1 and X_2 into quotient and remainder when divided by $\sqrt{n}$. These different components help us communicate the message. Let

$$X_1 = \lfloor\sqrt{n}\rfloor X_{1q} + X_{1r} \qquad (2.101)$$
$$X_2 = \lfloor\sqrt{n}\rfloor X_{2q} + X_{2r} \qquad (2.102)$$

where $0 \leq X_{1r}, X_{2r} \leq \lfloor\sqrt{n}\rfloor - 1$.

The intuition here is that for X_1 and X_1 to both have an influence on Y_1 and Y_2 they need to be not too far apart in value. Define

$$A = \begin{cases} 1 \text{ if } X_{1q} \geq X_{2q} + 2 \\ 0 \text{ if } X_{1q} - X_{2q} = 1 \text{ or } 0 \text{ or } -1 \\ -1 \text{ if } X_{1q} \leq X_{2q} - 2 \end{cases} \tag{2.103}$$

Then we have:

$$\begin{aligned}
&I(X_1; Y_1, Y_2|X_2) && (2.104)\\
&\leq I(X_1, A; Y_1, Y_2|X_2) && (2.105)\\
&= I(A; Y_1, Y_2|X_2) + I(X_1; Y_1, Y_2|X_2, A) && (2.106)\\
&\leq \log 3 + I(X_1; Y_1, Y_2|X_2, A) && (2.107)\\
&= \log 3 + I(X_1; Y_1, Y_2|X_2, A = 1)P(A = 1) \\
&\quad + I(X_1; Y_1, Y_2|X_2, A = 0)P(A = 0) + I(X_1; Y_1, Y_2|X_2, A = -1)P(A = -1) && (2.108)
\end{aligned}$$

Similarly, we can also show:

$$\begin{aligned}
&I(X_2; Y_1, Y_2|X_1) && (2.109)\\
&= \log 3 + I(X_2; Y_1, Y_2|X_1, A = 1)P(A = 1) \\
&\quad + I(X_2; Y_1, Y_2|X_1, A = 0)P(A = 0) + I(X_2; Y_1, Y_2|X_1, A = -1)P(A = -1) && (2.110)
\end{aligned}$$

First, consider the case $A = 1$. Then we have $X_{1q} \geq X_{2q} + 2$ and so

$$\begin{aligned}
X_1 - \lfloor\sqrt{n}\rfloor &= \lfloor\sqrt{n}\rfloor X_{1q} + X_{1r} - \lfloor\sqrt{n}\rfloor && (2.111)\\
&\geq \lfloor\sqrt{n}\rfloor(X_{1q} - 1) && (2.112)\\
&\geq \lfloor\sqrt{n}\rfloor(X_{2q} + 1) && (2.113)\\
&\geq X_2 && (2.114)
\end{aligned}$$

Therefore, $Y_1 = Y_2 = X_2$ if $A = 1$.

$$I(X_1; Y_1, Y_2|X_2, A = 1) = I(X_1; X_2, X_2|X_2, A = 1) = 0 \tag{2.115}$$

In this case both Y_1 and Y_2 only depend on X_1, and we can't get the benefit of having two links. The same thing holds if $X_{1q} \leq X_{2q} - 2$, i.e $A = -1$.

$$I(X_2; Y_1, Y_2|X_1, A = -1) = 0 \tag{2.116}$$

Further, in these cases we are bottlenecked by the entropy of one of the signals.

$$\begin{aligned}
I(X_1, Y_1, Y_2|X_2, A = 1) &\leq H(X_1) = \log(n + 1) && (2.117)\\
I(X_2, Y_1, Y_2|X_1, A = -1) &\leq H(X_2) = \log(n + 1) && (2.118)
\end{aligned}$$

The interesting case is when $A = 0$

$$I(X_1; Y_1, Y_2|X_2, A = 0) \tag{2.119}$$
$$= H(X_1|X_2, A = 0) - H(X_1|Y_1, Y_2, X_2, A = 0) \tag{2.120}$$
$$\le H(X_1|X_2, A = 0) \tag{2.121}$$
$$\le \log(3\lfloor\sqrt{n}\rfloor) \tag{2.122}$$

(For every choice of X_2 there are 3 choices for X_{1q} (since $A = 0$) and $\lfloor\sqrt{n}\rfloor$ choices for X_{1r}. Thus there are $3\lfloor\sqrt{n}\rfloor$ choices for X_1 given X_2 and $A = 0$)

$$\le \log 3 + \frac{1}{2}\log(n+1) \tag{2.123}$$

By symmetry we can show

$$I(X_2; Y_1, Y_2|X_1, A = 0) \le \log 3 + \frac{1}{2}\log(n+1) \tag{2.124}$$

Collectively, these give us

$$\max_{p(X_1,X_2)} \min\{I(X_1; Y_1, Y_2|X_2), I(X_2; Y_1, Y_2|X_1)\} \le \log 3 + \frac{1}{2}\log(n+1) + \log 3 \tag{2.125}$$
$$= 2\log 3 + \frac{1}{2}\log(n+1) \tag{2.126}$$

■

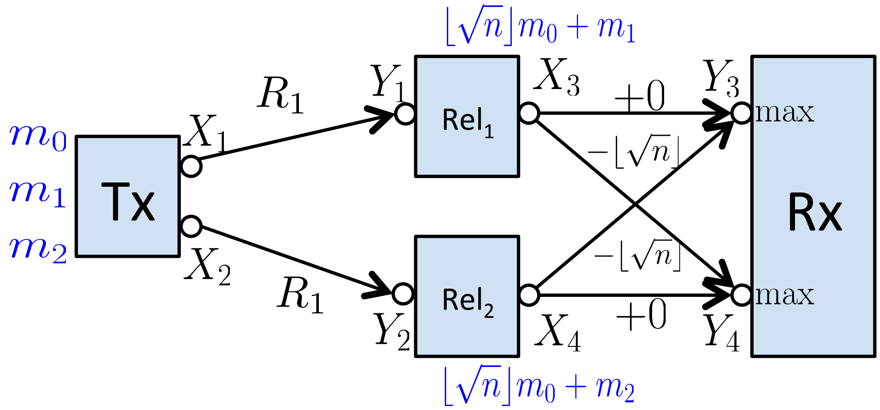


Figure 2.9: A max-superposition relay network with an approximately optimal partial-decode-and-forward scheme. The cross-links between X_3, X_4 and Y_3, Y_4 are weaker than the direct links and lead to downward bit-level shifts by $\lfloor\sqrt{n}\rfloor$. $X_1 = \lfloor\sqrt{n}\rfloor m_0 + m_1$ and $X_2 = \lfloor\sqrt{n}\rfloor m_0 + m_2$ are transmitted by the two relays.

Theorem 2.5.2. *The cutset bound applied to the system in Fig. 2.9 gives the following upperbound on the achievable rate R for the system. R_1 is the rate across the links between the transmitter and each of the relays.*

$$R \leq 2R_1 \tag{2.127}$$

$$R \leq R_1 + \frac{1}{2}\log(n+1) + 2\log 3 \tag{2.128}$$

$$R \leq \frac{3}{2}\log(n+1) + \frac{3}{2}\log 2 \tag{2.129}$$

Proof: Applying the cutset bound to (1) the cut between the transmitter and the two relays (eq. (2.130)), (2) the cut between the receiver and Rel_2 on one side and Rel_1 and the transmitter on the other side and the other symmetric cut (eq. (2.131)) and finally (3) the cut between the relays and the decoder (eq. (2.5.4)) gives:

$$R \leq 2R_1 \tag{2.130}$$

$$R \leq \max_{p(X_1,X_2)} \min\{R_1 + I(X_3;Y_2,Y_3,Y_4|X_4), R_1 + I(X_4;Y_1,Y_3,Y_4|X_3)\} \tag{2.131}$$

$$R \leq \max_{p(X_1,X_2)} I(X_3,X_4;Y_3,Y_4) \tag{2.132}$$

We get eq. (2.127) directly from eq. (2.130). To get eq. (2.128), we apply Lemma 2.5.1:

$$\max_{p(X_1,X_2)} \min\{R_1 + I(X_3;Y_3,Y_4|X_4), R_1 + I(X_4;Y_3,Y_4|X_3)\} \leq \frac{1}{2}\log(n+1) + 2\log 3 \tag{2.133}$$

Now, we investigate .

If $|X_3 - X_4| > \sqrt{n}$, then Y_3 and Y_4 both are equal to either X_3 or X_4. So $|\{(Y_3,Y_4)\}| \leq n+1$, and $I(X_3,X_4;Y_3,Y_4) \leq \log(n+1)$.

So to maximize the mutual information, the other case we want to consider is $-\lfloor\sqrt{n}\rfloor \leq X_3 - X_4 \leq \lfloor\sqrt{n}\rfloor$. In this case we have

$$-\lfloor\sqrt{n}\rfloor \leq Y_3 - Y_4 \leq \lfloor\sqrt{n}\rfloor \tag{2.134}$$

And then we have

$$|\{(Y_3,Y_4)\}| \tag{2.135}$$

$$\leq (n+1)\cdot(2\sqrt{n}+1) \tag{2.136}$$

(Once Y_3 is fixed ($(n+1)$ choices) there are at most $(2\sqrt{n}+1)$ choices for Y_4, such that $-\lfloor\sqrt{n}\rfloor \leq X_3 - X_4 \leq \lfloor\sqrt{n}\rfloor$)

$$\leq (n+1)\cdot\sqrt{8(n+1)} \tag{2.137}$$

$$= 2^{\frac{3}{2}}(n+1)^{\frac{3}{2}} \tag{2.138}$$

So we have

$$I(X_3,X_4;Y_3,Y_4) \leq H(Y_3,Y_4) = \frac{3}{2}\log(n+1) + \frac{3}{2}\log 2 \tag{2.139}$$

which gives us eq. (2.129). ∎

Chapter 3

Bit-level models for uncertain parameters: dynamical systems and active control

3.1 Introduction

This chapter extends the carry-free approach to model dynamical systems and active control. As expected of models that grew out of the information-theory community, it is very easy to model communication constraints in control systems using these models. The pictures associated with the model make it very easy to understand classic results in rate-limited control.

We build from the rate-limited control perspective to understand a carry-free uncertainty threshold principle. The uncertainty threshold principle is a classical result from control that provides a threshold on the tolerable uncertainty in system parameters for mean-squared stabilization. If the uncertainty is too high, the system cannot be stabilized. Carry-free models highlight the fact that the uncertainty threshold principle [6] stems from the same bottleneck that lead to the $\log \log SNR$ result for the capacity unknown fading channels [63] (discussed in Chapter 2.).

We start by introducing some basic notions of system stability in real-valued systems.

Definition 3.1.1. A system with state $X[n]$ at time n, where $X[n]$ is a random variable is said to be **mean-square stable** if there exists $M \in \mathbb{R}$, $M < \infty$, s.t. $\mathbb{E}[|X[n]|^2] < M$ for all n.

Definition 3.1.2. Let the evolution of system state $X[n]$ be governed by a linear system

$$\begin{aligned} X[n+1] &= A[n] \cdot X[n] + B[n] \cdot U[n] + W[n] \\ Y[n] &= C[n] \cdot X[n] + V[n] \end{aligned} \tag{3.1}$$

$A[n], B[n], C[n], V[n], W[n]$ are random variables with known distributions. Let $\sigma(Y_0^n)$ be the sigma-algebra generated by the random variables $Y[0]$ to $Y[n]$. The system is said to be **mean-square stabilizable** if there exists a causal control strategy $U_0^n(\cdot) \in \sigma(Y_0^n)$ such that $X[n]$ is mean-square stable, i.e. there exists $M \in \mathbb{R}$, $M < \infty$, s.t. $\mathbb{E}[|X[n]|^2] < M$ for all n.

Definition 3.1.3. Let the evolution of system state $X[n]$ be governed by a linear system

$$\begin{aligned} X[n+1] &= A[n] \cdot X[n] + W[n] \\ Y[n] &= C[n] \cdot X[n] + V[n] \end{aligned} \tag{3.2}$$

$A[n], B[n], C[n], V[n], W[n]$ are random variables with known distributions. The system is said to be **mean-square observable** if there exists a casual estimator $\widehat{X}[n]$ that is a function of $Y[1], Y[2], \cdots Y[n]$ such that $|X[n] - \widehat{X}[n]|$ is mean-square stable.

3.2 Rate-limited estimation and control

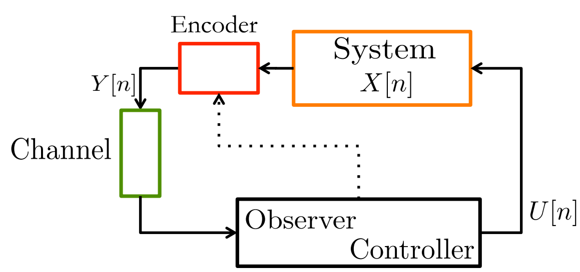


Figure 3.1: The rate-limited control allows the system to use an encoder and decoder for communication across the rate-limited sensing channel. We can assume the decoder box is merged with the observer/controller box and is not shown separately here. Further, the controller provides feedback of the control signal sent to the encoder box so that only the innovation in the system need be encoded.

Information theorists look to rate as the parameter that characterizes a communication system. As discussed in the introduction, a natural first attack is to bring in tools from information theory and look at control over rate-limited channels. This setup was first considered in papers such as [117, 79] and [110]. Consider,

$$X[n+1] = A \cdot X[n] + W[n] \ . \tag{3.3}$$

Let $W[n] \sim \mathcal{N}(0,1)$ i.i.d., and $X[0] \sim \mathcal{N}(0,1)$. A is a known constant. With a perfect observation, $Y[n] = X[n]$ this system is clearly mean-square observable.

In [110], Tatikonda and Mitter looked at the problem of first estimating and then controlling an unstable system over a noiseless rate-limited channel, with access to a full encoder and decoder on either end of the channel (Fig. 1.2). They showed that for finite estimation error the channel must be able to support a rate $R > \sum \log |\lambda(A)|$, where $\lambda(A)$ are the unstable eigenvalues of the matrix A. The proof technique was non-standard in control at the time, and hinged on showing that the volume of uncertainty in the system was growing as the product of the unstable eigenvalues of A.

$$X[n+1] = A \cdot X[n] + U[n] \tag{3.4}$$

The basic argument for the control problem in eq. (3.4) is based on the separation principle and follows from the estimation result. Since the controller has memory of previously applied controls, it can simulate an observer and estimate the system state in open-loop. Using this, it can also estimate the closed-loop state and compute the optimal control from the current state-estimate, as the one that will cancel out the state. The necessary condition on the rate is also given from the estimation results. Subtracting the state estimate at each time (using knowledge of the controls that have been previously applied by the controller as in Fig. 3.1) gives a matching sufficient condition. Full knowledge of the applied controls is critical in this strategy, and the results in [110, 96] use control feedback to the encoder to achieve stabilization.

This idea extends to other data-rate theorems for control over (noiseless and noisy) rate-limited channels, as in Sahai and Mitter [96], Nair et al. [79], and more recently in results like [75]. In these problems, it is common to think of the observer and controller as lumped into one box (Fig. 3.1), with a noiseless channel between controller and actuator.

3.2.1 Bit-level models

ADT models

We can use ADT models [7] to understand the information bottlenecks in the real-valued systems described in the earlier section. Let us consider a scalar system, with $A = a$. Then, building from the previous chapter, let the state as it evolves in time be represented in binary expansion by the bits

$$x_m[n]x_{m-1}[n] \cdots x_1[n]x_0[n].x_{-1}[n]x_{-2}[n] \cdots . \tag{3.5}$$

For convenience, we'll call this sequence of bits the state at time n. m is the highest non-zero bit-level occupied by the state at time n. In formal polynomial notation, we denote the state as

$$x[n](z) = x_m[n]z^m + x_{m-1}[n]z^{m-1} + \cdots + x_0[n] + x_{-1}z^{-1} \cdots . \tag{3.6}$$

The system noise is $w[n](z) = w_{-1}[n]z^{-1} + w_{-2}[n]z^{-1} + \cdots$. Thus, all bits below the zero level are replaced at each time step. The gain of the system is fixed as $a[n](z) = a(z) = 1 \cdot z^{g_a}$.

Then, we can think of the open-loop state evolution as

$$x[n+1](z) = a(z)x[n](z) + w[n](z) \tag{3.7}$$

Intuitively we can see that the stack of bits is growing taller by g_a bits with each clock tick (Fig. 3.2) as discussed in the introduction to the dissertation. The relationship between the real-valued scalar gain a, and the ADT gain $a(z)$ is through the degree of $a(z)$. Since the ADT models represent numbers in binary, we have $g_a = \log a$. Hence, the minimum communication rate required for estimation is given by the number of new bits that are pulled into the system state at each time, which is $\log a$.

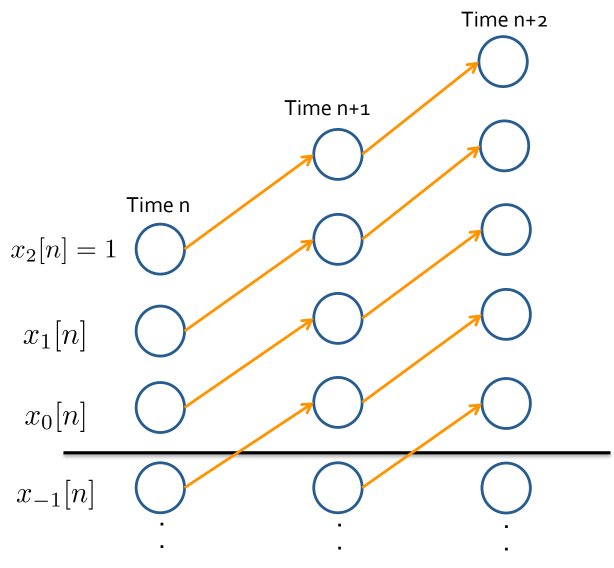


Figure 3.2: The system state can be thought of as a stack of bits marching upward with the gain a.

For control, we can consider the closed-loop system

$$x[n+1](z) = a(z)x[n](z) + u[n](z) + w[n](z) \tag{3.8}$$

where $u[n](z)$ is the control applied at time n.

The control system needs the ability to dissipate the new uncertainty in the system pulled into the system. To dissipate g_a bits, it must know about g_a of them. So the same rate-threshold applies for stabilizability with control: for $R > g_a$ we can ensure that the magnitude of the system is bounded.

We can think of these bit-level models as an alternative perspective to quantization-box or mutual-information style understandings of information flow for these well known problems. With deterministic system gains we see that the underlying uncertainty is initialized by the initial state $x[0]$, and then is refreshed at each time. The unstable gain pumps an additional g bits from that reservoir above the decimal point at each time and creates the need to dissipate the uncertainty.

Carry-free models

We now express the same idea as a carry-free model to build towards the next section. Both ADT models and carry-free models are particularly conducive to modeling explicit rate constraints. A rate limit simply caps the number of levels that are visible to the estimator or controller at any given time.

The carry-free model for this particular system with a fixed gain $a(z)$ is essentially identical to the ADT model, except now we no longer restrict $a(z)$ to be a monomial. $a(z) = 1 \cdot z^{g_a} + a_{g_a-1}z^{g_a-1} + \cdots$, and is not time varying. The highest bit of the gain $a[n](z)$ is always 1. Fig. 3.3 shows a bubble-picture for this system. This figure captures the same effect as Fig. 1.4 except that the multiplication has a different effect.

Thus, the state evolution is still as below, but we use carry-free addition and multiplication.

$$x[n+1](z) = a(z)x[n](z) + w[n](z) \tag{3.9}$$

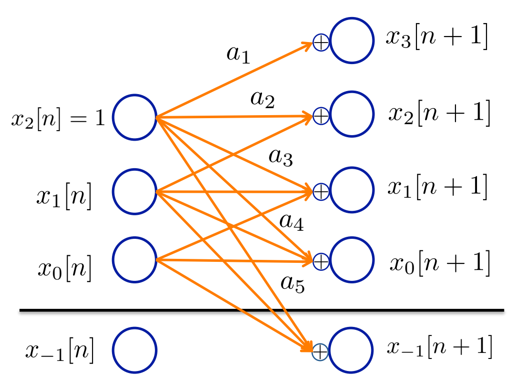


Figure 3.3: Carry-free for models with highest bit at level 3, and power $g = 1$.

In this case, again the uncertainty is refreshed by the randomness introduced by the additive system noise. At each time step, the gain g_a pulls up g_a new bits of information from the noise and introduces it into the state above the noise level. Then if $R > g_a$ (i.e. our observation allows us to infer the same volume of information as is being introduced into the system) we will be able to keep track of the state even though it grows at every time step. The control problem behaves similarly, since we assume we have a perfect reliable control channel.

3.3 The uncertainty threshold principle

The uncertainty threshold principle [6] is a classical result from the control literature that deals with parameter uncertainty in systems. It provides a threshold beyond which the uncertainty in the system is simply too large for the system to be stabilized. Consider the system below:

$$\begin{aligned} X[n+1] &= A[n] \cdot X[n] + B[n] \cdot U[n] + W[n], \\ Y[n] &= X[n] \end{aligned} \tag{3.10}$$

Both $A[n]$ and $B[n]$ are Gaussian i.i.d. sequences. Let $\mathbb{E}[A[n]] = \mu_A, Var(A[n]) = \sigma_A^2$, and $\mathbb{E}[B[n]] = \mu_B, Var(B[n]) = \sigma_B^2$. $A[n], B[m]$ are independent for all n, m and are also independent of all system parameters[1]. $\mathbb{E}[W[n]] = 0, Var(A[n]) = 1$, i.i.d..

Theorem 3.3.1. *(from [6]) The system in (3.10) can be stabilized in a mean-squared sense if and only if* $\sigma_A^2 + \mu_A^2 - \frac{(\mu_A \mu_B)^2}{\sigma_B^2 + \mu_B^2} < 1$.

This condition is derived using a standard stochastic dynamic programming approach using a Riccati equation. However, these techniques hide the flow of information in the system that actually lead to these bottlenecks. We will use carry-free models to provide an information-centric perspective on this.

To better understand the information bottlenecks in this system we separately consider randomness first on the system gain $A[n]$ and then on the control $B[n]$ so we can understand how they behave differently.

If $B[n]$ is a constant this threshold reduces to $\sigma_A^2 < 1$. If $A[n] = \mu_A$ (a constant) the condition becomes $\frac{\sigma_B^2 \mu_A^2}{\sigma_B^2 + \mu_B^2} < 1$. We will explore these particular setups in detail using a carry-free model below.

3.3.1 Unknown system gain

For systems with random system gain, we first consider the simplest problem of predicting a system that is evolving as below (3.11). This is just to highlight that the duality between observation and control breaks down in the presence of random parameters. Hence, traditional techniques might not work to incorporate rate limitations for these problems and the subsequent carry-free techniques we introduce will be particularly valuable.

$$\begin{aligned} X[n+1] &= A[n] \cdot X[n] + W[n], \\ Y[n] &= X[n] \end{aligned} \tag{3.11}$$

[1] Athans and Ku [6] also allow for a cross correlation term between $A[n]$ and $B[n]$ and provide a threshold for this case. We focus on the case where all the random parameters are independent.

Here, $A[n] \in \mathbb{R}$ can randomly vary with time, and $W[n]$ is additive noise. The observer observes $X[n]$ perfectly. Let $A[n] \sim \mathcal{N}(\mu_A, \sigma_A^2)\ \forall\ n$ be i.i.d.. $w[n] \sim \mathcal{N}(0,1)$ also i.i.d. and independent of other system parameters. Let $Y[n] = X[n]$, i.e. full past state information is available to the estimator/controller. Let $\widehat{X}[n+1]$ be state-estimate by the estimator at time n, and let the expected estimation error $\mathbb{E}[e[n]^2] = \mathbb{E}(X[n+1] - \widehat{X}[n+1])^2$.

Theorem 3.3.2. *The system (3.11) is not mean-square predictable if $\mu_A > 1$.*
Proof:

$$\widehat{X}[n+1] = \mathbb{E}[X[n+1] \mid Y[n]] \tag{3.12}$$
$$= \mathbb{E}[A[n]]X[n] + \mathbb{E}[W[n]] = \mu_A X[n] \tag{3.13}$$

$$\mathbb{E}[e[n]^2] = \mathbb{E}[(A[n] - \mu_A)^2 X[n]^2] + \mathbb{E}[W[n]^2] \tag{3.14}$$
$$= \sigma_A^2 \mathbb{E}[X[n]^2] + 1 \tag{3.15}$$

where (3.14) follows since $W[n]$ is independent of the state $X[n]$. Hence, our estimation error grows as the state $\mathbb{E}[X[n]^2]$, which tends to infinity as $n \to \infty$ if $\mu_A > 1$. ∎

This shows that prediction can be a hard problem if the gain of the system is larger than 1 and varying. Even with perfect observation of the current state, there is not much we can do.

It might seem that control for such systems is a lost cause given that prediction is impossible. However, we know from the uncertainty threshold principle that in fact such systems are stabilizable if $\sigma_A^2 < 1$. It is interesting to note here that μ_A does not play any role whatsoever. The prediction error thus does not depend on how fast the system is growing, but only the amount of uncertainty introduced by the system gain. We state this theorem to be able to later compare it to the corresponding carry-free model.

Consider the system, where we have access to a control signal as in (3.16).

$$\begin{aligned} X[n+1] &= A[n]X[n] + U[n] + W[n], \\ Y[n] &= X[n] \end{aligned} \tag{3.16}$$

$a[n] \sim \mathcal{N}(\mu_A, \sigma_A^2)\ \forall\ n$ be i.i.d.. $W[n] \sim \mathcal{N}(0,1)$ also i.i.d. Instead of just estimating the system state, the task is to stabilize the system.

Theorem 3.3.3. *System (3.16) is stabilizable in a mean-square sense if $\sigma_A < 1$.*
Proof: $(\Longleftarrow)$
Choose $U[n] = -\mu_A \cdot Y[n] = -\mu_A \cdot X[n]$. Then,

$$\mathbb{E}[X[n+1]^2] = \mathbb{E}[(A[n] - \mu_A)^2 X[n]^2] + \mathbb{E}[W[n]^2] \tag{3.17}$$
$$= \sigma_A^2 \mathbb{E}[X[n]^2] + 1 \tag{3.18}$$

For such a recurrence, $\mathbb{E}[X[n+1]^2]$ is finite for all $\sigma_A < 1$. ∎

So we can control a state of an unstable system ($\mathbb{E}[A[n]] > 1$), even though we would not able to predict it if it were uncontrolled.

The following sections will use carry-free models to illustrate the underlying information flow in these systems that lead to the difference in the prediction/estimation and the control cases.

3.4 Carry-free models for the uncertainty threshold principle

The carry-free models in Chapter. 2 show that parameter uncertainty is essentially also an informational bottleneck that plays a role similar to a rate limitation, even though there is no explicit rate constraint. We know this from Lapidoth's [63] $\log\log SNR$ result, the intuition for which was seen in the previous chapter. As the SNR goes to infinity, so does the capacity. But all bits much lower than the MSB are lost, as we saw earlier. Reaching capacity requires a careful optimization on the input distribution.

Carry-free models therefore allow us to merge explicit rate constraints as considered earlier in this chapter, with parameter uncertainty and multiplicative noise. Using these models we try to give a unifying perspective on the information flows in systems, be they because of rate constraints, additive noise or multiplicative noise. One of the most interesting points that we will show in this section are that in fact, the informational bottleneck that leads to the uncertainty threshold principle is the *same one* that leads to the $\log\log SNR$ result.

Let us now develop these carry-free models for these systems.

The state as it evolves in time is represented by the bits

$$x_{d_n}[n]x_{d_n-1}[n]\cdots x_1[n]x_0[n].x_{-1}[n]x_{-2}[n]\cdots, \tag{3.19}$$

and we use polynomial notation as below:

$$x[n](z) = x_{d_n}[n]z^{d_n} + x_{d_n-1}[n]z^{d_n-1} + \cdots + x_0[n] + x_{-1}z^{-1}\cdots. \tag{3.20}$$

d_n is the degree of the polynomial at time n. For ease of notation, we will often suppress the dummy polynomial variable z wherever obvious. The additive system noise is $w[n](z) = w_{-1}[n]z^{-1} + w_{-2}[n]z^{-1} + \cdots$. Thus, all bits below the zero level are replaced at each time step as in the ADT model.

3.4.1 Estimation with unknown system gain

First we consider the carry-free parallel to the eq. (3.11). Let $a[n](z) = 1\cdot z^{g_a} + a_{g_a-1}[n]z^{g_a-1} + a_{g_a-2}[n]z^{g_a-2} + \cdots)$, with the first bit fixed as 1. Thus the system has one deterministic bit. The restriction of the first bit of $a[n](z)$ to be 1 just ensures that there is no uncertainty in the position of the highest bit, even though lower down individual bits may be random. We can think of this as having knowledge of the variance of a random gain in the real system. This idea will be generalized to finer models later.

Beyond the first bit, each $a_i[n]$ is a Bernoulli-($\frac{1}{2}$) random variable. The observer noiselessly observes the entire state $x[n]$. This system evolves according to the equation

$$\begin{aligned} x[n+1](z) &= a[n](z) \cdot x[n](z) + w[n](z), \\ y[n](z) &= x[n](z) \end{aligned} \tag{3.21}$$

The state evolution in this system can be thought of as a carry-free channel with unknown fading [85], with channel gain $a[n]$, channel input $x[n]$ and output $x[n+1]$. Figure 3.4 shows the bit-level interactions.

We define the estimation error for this problem.

Definition 3.4.1. Let $\widehat{x}[n](z)$ be any estimate of $x[n](z)$ based on $y[n]$. Then the estimation error at time n is defined as $e[n](z) = x[n](z) - \widehat{x}[n](z)$.

Theorem 3.4.2. *Let $g_e[n]$ be the highest non-zero bit-level of $e[n](z)$. If the system gain, $g \geq 1$, then $\mathbb{E}[g_e[n]] \to \infty$, i.e. the degree of the error polynomial is unbounded.*

Proof: Recall d_n is the degree of $x[n]$ and thus represents the position of the highest non-zero bit. We know from Lemma 2.4.1 in the previous chapter that $I(x[n]; x[n+1]) = I(d_n; d_{n+1})$, i.e. the position of the maximal non-zero bits are sufficient statistics. This means that the observation $x[n]$ only contains information about the maximal bit of $x[n+1]$. All of the remaining information has been lost due to the scrambling by the random bits of $a[n]$. For example, in Fig. 3.4, with random a_i, we are only able to predict $x_3[n+1]$.

Further, note that d_{n+1} is a deterministic function of d_n, i.e. $d_{n+1} = d_n + g_a$. Hence, the observations after $y[0]$ do not provide any new information. With $g_a > 1$ then $\mathbb{E}[d_n] \to \infty$ as $n \to \infty$. Hence, $\mathbb{E}[d_n - 1]$ also goes to infinity. Since x_{d_n-1} cannot be predicted, $\mathbb{E}[g_e[n]] \to \infty$. ■

3.4.2 Control with unknown system gain

What changes in the case of control? The carry-free parallel to system (3.16) is:

$$\begin{aligned} x[n+1](z) &= a[n](z) \cdot x[n](z) + u[n](z) + w[n](z), \\ y[n](z) &= x[n](z) \end{aligned} \tag{3.22}$$

We define a notion of stability for carry-free control systems that is similar to "in-the-box" stability. The system is stabilizable if it is always bounded.

Definition 3.4.3. Recall that d_n is the degree of $x[n]$ at time n. Then the system (3.22) is said to be stabilizable if there exists $u[n]$ that is a function of $y[1], \cdots y[n]$, and $M < \infty$, such that $d_n < M$ for every n.

Theorem 3.4.4. *System.* (3.22) *is stabilizable if $g_a \leq 1$.*

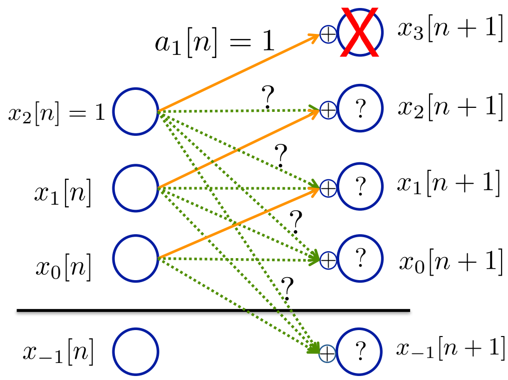


Figure 3.4: An observer can only see the top bit $x_3[n+1]$. However, a controller can act and kill $x_3[n+1]$ to maintain the level of the state at 3 bits above the noise level.

Proof: ($\Longleftarrow$)
Let $a[n] \cdot x[n] = p[n]$. Again, d_n is the degree of $x[n]$. The degree of $a[n]$ is g_a, and let the degree of $p[n]$ be $g_{p,n}$, and the degree of $y[n]$ be $g_{y,n}$.

This proof builds off of Corollary 2.4.1, which implies that: $I(p[n]; u[n]) = I(p[n]; y[n]) = I(g_{p,n}; g_{y,n}) = I(g_{p,n}; d_n)$.

Now, since the highest bit of $a[n]$, $a_g[n] = 1$, the controller can exactly know the position of the highest bit of $x[n]$ based on $y[n]$. So d_n is perfectly known to the controller. So the controller can make sure it chooses a control that cancels the first bit of $x[n]$. It chooses $u[n](z) = z^{d_n + g_a}$. The controller can therefore always maintain $d_{n+1} = d_n$, as long as the increase in uncertainty through g_a is no more than the deterministic bits in the gain, which is 1. So if $g_a \leq 1$, the system is stabilizable. See Fig. 3.4.

∎

The interesting phenomenon here is that we can reduce the uncertainty about the system by changing the system state. The carry-free pictures provide an illustration for this and an intuition for what leads to the difference between estimation and control.

3.4.3 Unknown control gain

Now consider the carry-free system with a random gain for the control input, but constant gain on the system:

$$x[n+1](z) = a(z) \cdot x[n](z) + b[n](z) \cdot u[n](z) + w[n](z) \tag{3.23}$$

We restrict attention to the case where the gain on the state is a known constant $a(z) = 1 \cdot z^{g_a}$ for all n.

We model the random bits of the gain $b[n](z)$ as Bernoulli($\frac{1}{2}$) variables. g_b is the highest non-zero bit-level of $b[n](z)$ for every n. The model in [85] focused on random gains where all the bits were random. However, this model does not accurately capture non-zero-mean random variables. We generalize that model to the one in eq. (3.25), which allows for both deterministic and random bits. Without loss of generality, we can assume that the deterministic bits of the gain $b[n]$ are a 1 followed by a series of 0's. All of our arguments would extend to any other set of deterministic bits, led by a 1.

$$b[n](z) = b_{g_b}[n] \cdot z^{g_b} + b_{g_b-1}[n] \cdot z^{g_b-1} + \cdots \tag{3.24}$$

$$= 1 \cdot z^{g_{det}} + 0 \cdot z^{g_{det}-1} + 0 \cdot z^{g_{det}-2} + \cdots + b_{g_{ran}}[n] \cdot z^{g_{ran}} + b_{g_{ran}-1}[n] \cdot z^{g_{ran}-1} + \cdots \tag{3.25}$$

Define:

$$g_{ran} = \max\{i | b_i \sim \text{Bernoulli}(\frac{1}{2}), g_b \geq i \geq -\infty\}. \tag{3.26}$$

g_{ran} is the index of the highest random bit of $b[n](z)$ $\forall n$. g_{ran} does not vary with n. On the other hand, the value of the bit at level g_{ran}, i.e. $b_{g_{ran}}[n]$, does. Further, $b_i[n] \sim \text{Bernoulli}(\frac{1}{2}), \forall i \leq g_{ran}$ and the realizations of these bits are unknown to the controller. If $g_{ran} < g_b$, then we also define g_{det} as the highest deterministic bit-level of b. $b_{g_{det}} = 1$. All bits from $b_{g_{det}-1}$ to $c_{g_{ran}+1}$ are fixed to be 0, and these bits are known to the controller [2].

$$g_{det} = \begin{cases} \max\{i | b_i = 1\} \text{ if } g_{ran} < g_b \\ g_{ran}, \text{ otherwise.} \end{cases} \tag{3.27}$$

By construction, we have $g_{det} \geq g_{ran}$. So there are $g_{det} - g_{ran}$ deterministic bits in the gain, and all lower bits are random. See Fig. 3.5 for details.

Theorem 3.4.5. *The carry-free system (3.23) is stabilizable if and only if $g_{det} - g_{ran} \geq g_a$ for $g_a > 0$. For $g_a \leq 0$ the system is self stabilizing.* This system is extensively discussed in Chapter 5, and so here we provide an intuition.

Intuition At every time step the magnitude of the system state increases by g_a bits. $g_{det} - g_{ran} \geq g_a$ is sufficient to ensure that the state does not grow in magnitude. Further, the unknown nature of the control only allows us to control only $g_{det} - g_{ran}$ bits of the state — the controller can force the top $g_{det} - g_{ran}$ bits of the state to any value it desires (since it knows what the state there is going to be), but the next bit can be correct with probability only $\frac{1}{2}$. So with probability $\frac{1}{2}$ there will be an increase in the degree of $x[n]$, i.e. d_n, which we cannot tolerate. See Fig. 3.5. ■

[2] The bits are fixed to zero for simplicity. This can be done without loss of generality, since we can choose $u[n]$ appropriately to cancel the right bits of $x[n]$ as long as $b_{g_{det}} = 1$. The values of the bits of $u[n]$ would have to be set differently for different values of the deterministic bits by solving the linear equations that related them to the bits of $x[n]$.

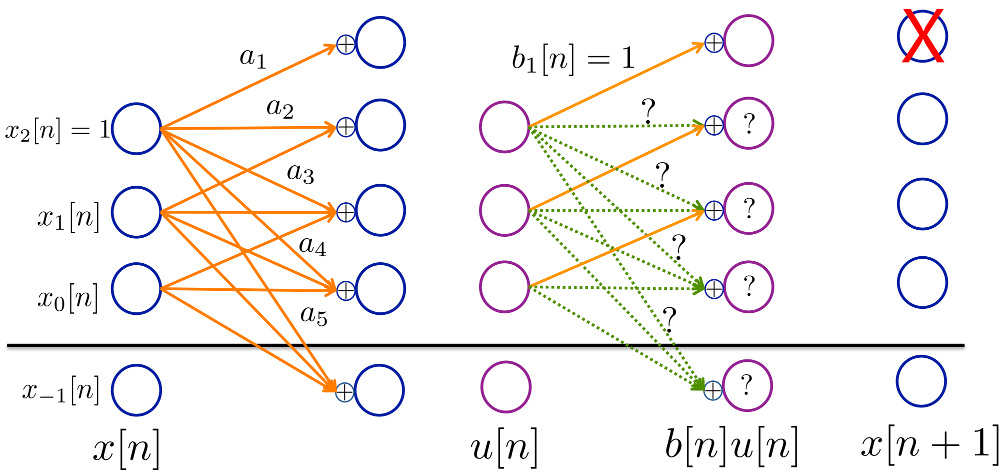


Figure 3.5: Carry-free model for system (3.23). The system gain has $g_a = 1$, $g_{det} = 1$ and $g_{ran} = 0$. So $b_1[n] = 1$ is a deterministic value that is the same at each time step. $g_{ran}[n] = 1$, so bits $b_0[n], b_{-1}[n], b_{-2}[n], \cdots$ are all random Bernoulli-($\frac{1}{2}$) random bits. As a result the controller can only influence the top bit of the control going in.

Consider the uncertainty threshold principle for the corresponding real-valued system $\frac{\sigma_B^2 \mu_A^2}{\sigma_B^2 + \mu_B^2} < 1$ with constant system gain μ_A. If we take logs on both sides this condition is the same as $\log \sigma_B^2 - \log(\mu_B^2 + \sigma_B^2) + \log \mu_A^2 < 0$ or $\log(\mu_B^2 + \sigma_B^2) - \log \sigma_B^2 > \log \mu_A^2$.

$\log(\mu_B^2 + \sigma_B^2)$ is the counterpart to g_{det} and $\log \sigma_B^2$ is the counterpart to g_{ran}. System growth g_a is the counterpart to $\log \mu_A^2$. So as long as the number of bits we can control is larger than the growth rate of the system, we can stabilize. This matches with the prediction of the carry-free model, considering that the carry-free model is a bit-level representation.

Remark 3.4.6. Unlike the case with deterministic system gains, ADT models would not suffice to understand systems with random control gains, since they only capture bit shifts. The loss of information due to multiplicative scrambling by the random gains is essential to understand the uncertainty bottleneck.

Remark 3.4.7. The bit-levels in the state fall into different categories. First, some bits are simply unpredictable since they are hit by the random part of the system gain or the random additive system disturbance. The second category of bits are predictable and these we hope to control. Our control is limited to $g_{det} - g_{ran}$ bits of those that are predictable. In this sense, the "shakiness" of the control limits the number of bits that we can reliably alter the same way that observation noise prevents us from learning anything about certain bits in the state. However, there is a crucial difference between the observation and control cases. As soon as we aim to control the bits that we can control, we transmute a whole set of bits that might have been known into being unknown, since they are beyond the $g_{det} - g_{ran}$ limit imposed by the random control gain.

Chapter 4

Non-coherent sensing channels

4.1 Introduction

This chapter explores control and estimation over non-coherent sensing channels, i.e. channels with multiplicative noise.

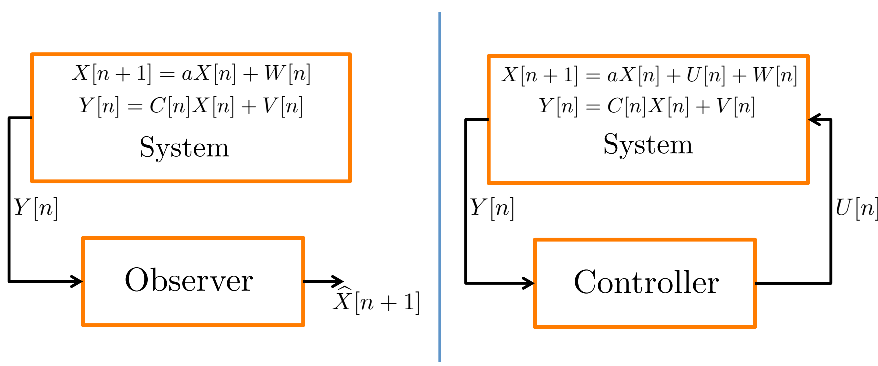


Figure 4.1: This chapter considers the estimation and control of a system with random multiplicative observation gain, $C[n]$, and random initial state $X[0]$. $V[n], W[n]$ are white noise and a is a scalar constant.

Non-coherent channels in communication systems emerge when phase-noise, frequency hopping or fast-fading make it impossible for the receiver to perfectly track the channel state. Such communication has been studied as far back as 1969 when Richters conjectured that even for continuous unknown fading-distributions the optimal input distribution is discrete [94]. The conjecture was proved in 2001 [1, 59]. Since then we know that the capacity of channels with unknown Gaussian fading scales as $\log\log(SNR)$ as opposed to the $\log(SNR)$ scaling of channels with known fading [108, 63].

Systems with multiplicative noise were explored around the same time as questions regarding non-coherent transmission were being asked. In 1971, Rajasekaran et al. derived the optimal *linear* filter for systems with unknown continuous observation gain [91]. However, this work did not consider non-linear strategies. This optimal linear filter was further analyzed by Tugnait [112], who noted that error turns out to be stable only in situations when the system is *stable*. The question of whether a non-linear filter could do better remained open. Kalman filtering with multiplicative noise was studied in [121, 42, 40], however these works were also limited to *a priori* stable systems.

The interaction of information theory and control theory has highlighted the importance of non-linear strategies (e.g. [115], [48]), and it is natural to explore the more general setting for the problems above. The optimality of discrete input distributions for transmission over non-coherent channels suggests that optimal estimation and control strategies for non-coherent systems may also be non-linear. However, to the best of our knowledge, these have not been investigated.

In fact, the recently introduced MMSE dimension [120] can be interpreted as characterizing the gain achievable by non-linear estimation over linear estimation in the high-SNR regime. The MMSE dimension measures the scaling behavior of the MMSE relative to the SNR as SNR grows in the presence of *additive* noise. For purely continuous distributions, the MMSE-dimension is 1, which implies that non-linear estimation offers the same scaling as linear estimation. This leads to a conjecture for the case of multiplicative noise: for systems with continuous distributions on the state, linear strategies should be essentially optimal. On the other hand, the MMSE-dimension for discrete random variables is 0, i.e. non-linear estimation can offer significant gains. In a control problem, actions may or may not be able to discretize a continuous-valued random state[1]. So can we use control to estimate better?

Unfortunately, the results from [120] cannot be directly applied to the problem at hand since the noise is multiplicative, and a naive strategy of taking logs of the system also does not help as is discussed later in Sec. 4.4.2 (Footnote 6). Instead, we reduce the estimation problem to a hypothesis-testing problem and use that to provide a lower-bound for the open-loop estimation error that itself must grow unboundedly with time.

The estimation result in this chapter contrasts with previous works in this vein that provide insightful thresholds to serve as design guidelines for systems. These thresholds provide a metric for the "uncertainty" in the system and provide regimes for observability and controllability. For instance, intermittent Kalman filtering results provide a threshold for the observation packet-drop probability parameter p, below which it is possible to track the system in a mean-square sense [102, 77, 86]. Neither the intermittent Kalman filtering setup nor the setup in this chapter allow for coding over the observations, however, the discrete randomness in the Kalman filtering result allows for a threshold result.

Tatikonda and Mitter provide a joint threshold for observability and controllability with R bits of feedback with coding allowed [110]. $R > \log a$, where $a > 0$ is the gain of a scalar, linear system is necessary and sufficient for both estimation and control. The separation principle states that often the optimal control is a function purely of the optimal state estimate. This leads to the *estimate-then-control* paradigm. Tatikonda and Mitter in [110]

[1]In this problem, it turns out they cannot.

observe: "If the state estimation error increases with time in an unbounded fashion there will come a point when we can no longer satisfy the control objective."

This raises the natural question: can a system that cannot be tracked in open-loop be stabilized with feedback? The answer is: sometimes yes. The ability to interact with the system lets the controller use a simple linear strategy stabilize if the uncertainty is small enough. This is reminiscent of the uncertainty threshold principle [6], which gives stabilizability threshold in the presence of random multiplicative system gains with perfect noiseless observations[2].

The control problem here is similar to the larger set of rate-limited control and estimation problems such as [30], which looks at control over fading channels with erasures, or [81, 75] which provide a data-rate threshold for stabilization. Even though technically the feedback in the problem at hand may include infinitely many bits, the systems's inability to choose and code these bits precludes the transmission of useful information.

So why are non-coherent observations different? One intuition is as follows: since the system is unstable, even a small multiplicative observation uncertainty leads to a large variation in the state estimate. The proof of the negative estimation result comes from first ignoring the additive noise in the state evolution and observations. Then the estimation problem is reduced to a hypothesis-testing problem with the help of a carefully constructed genie who sidesteps the prior on the initial state to provide side-information to the observer. With this reduction, the Chernoff-Stein lemma provides a lower-bound on the MMSE. The key technical lemmas for estimation are presented in Sec. 4.4.1 to distill the components above, before the proof of the theorem in given in Sec. 4.4.2.

We start by setting up the problem below (Sec. 4.2). Sec. 4.3 uses carry-free models to provide an intuition for why the estimation error cannot be tracked and illustrates the mechanics of a control strategy[3].

4.2 Problem setup

Before we set up the problem, some notational conventions: scalars and constants are denoted by lower-case, random variables are denoted by upper-case letters, realizations are lower-case. We use the Y_0^n to denote the sequence $Y[0], Y[1], \cdots Y[n]$, and so on. The density of X is denoted by $f_X(\cdot)$, and expectation by $\mathbb{E}[X]$.

[2] The estimation results in this chapter can also be thought of as the observability counterpart to the uncertainty threshold principle.

[3] As a side note, [120] implies that a finite number of bits of side information cannot change the scaling behavior of the MMSE, and carry-free models provide a visual interpretation for this.

4.2.1 Estimation

Consider the real-valued discrete-time scalar linear system (4.1) as shown in Fig. 4.1:

$$\begin{aligned} X[n+1] &= a \cdot X[n] + W[n], \\ Y[n] &= C[n] \cdot X[n] + V[n] \end{aligned} \tag{4.1}$$

$a \in \mathbb{R}, a > 1$ is a fixed and known scalar. $W[n]$ and $V[n]$ are i.i.d. white noise $\mathcal{N}(0,1)$ at each time n. $C[n]$ are i.i.d. random variables.

The observer has perfect recall and receives $Y[n]$ at time n. Let $\mathcal{F}_n$ be the σ-algebra defined by Y_0^n. The observer generates an estimate $\widehat{X}[n+1] \in \mathcal{F}_n$ of $X[n+1]$ at time n, based on all the past observations Y_0^n. Then the mean-squared estimation error is $\mathbb{E}[Err[n]^2] = \mathbb{E}[|X[n+1] - \widehat{X}[n+1]|^2]$. This chapter will focus on systems with only a multiplicative bottleneck and no additive noise for the estimation problem.

4.2.2 Control

Consider control of the same system:

$$\begin{aligned} X[n+1] &= a \cdot X[n] + U[n] + W[n], \\ Y[n] &= C[n] \cdot X[n] + V[n] \end{aligned} \tag{4.2}$$

where $C[n] \sim N(\mu, \sigma^2)$, μ, σ are known constants and $\sigma^2 \neq 0$. $X[0] \sim \mathcal{N}(0,1)$. Again, $\mathcal{F}_n$ is the σ-algebra defined by Y_0^n, and $U[n] \in \mathcal{F}_n$.

While we focus on scalar systems here, the results also extend to vector systems[4].

4.3 Carry-free models for non-coherent systems

Chapter 2 discussed the development of carry-free models to understand non-coherent communication. Here, we use these models to understand non-coherent estimation and control.

Let the state as it evolves in time be represented by bits as

$$x[n] = x_{d_n}[n] x_{d_n-1}[n] \cdots x_1[n] x_0[n] . x_{-1}[n] x_{-2}[n] \cdots \tag{4.3}$$

n is the time index. The subscript denotes the bit-level, i.e. d_n is the highest non-zero bit-level occupied by the state at time n. In polynomial notation, we denote the carry-free state as $x[n](z)$.

$$x[n](z) = x_{d_n}[n] z^{d_n} + x_{d_n-1}[n] z^{d_n-1} + \cdots + x_0[n] + x_{-1} z^{-1} \cdots \tag{4.4}$$

[4]For estimation, consider a genie that provides the exact values of all but one component of the vector valued state to the observer. The arguments here show that even that last remaining component cannot be estimated. For control, standard extensions from the literature (e.g. looking at the spectral radius of the system) would establish a sufficient condition for control.

We will suppress the dummy polynomial variable z wherever obvious. The gain of the system is fixed as $a(z) = 1 \cdot z^{g_a}$ and is not time varying.

For simplicity we ignore the driving disturbance. Define the state evolution as:

$$x[n+1] = a \cdot x[n] \tag{4.5}$$

The initial state $x[0](z)$ is a binary polynomial of degree d_0, with Bernoulli-$(\frac{1}{2})$ coefficients $P(x_i[0] = 1) = \frac{1}{2}$, $\forall\ i \leq d_0$. Eq. (4.5) implies that at each time step, the system magnitude increases by g_a bit-levels due to the gain $a(z)$. The observation $y[n](z)$ is:

$$y[n] = c[n] \cdot x[n] \tag{4.6}$$

We model the random bits of the gain $c[n](z)$ as Bernoulli$(\frac{1}{2})$ variables. g_c is the highest non-zero bit-level of $c[n](z)$ for every n. The model in [85] focused on random gains where all the bits were random. However, this model does not accurately capture non-zero-mean random variables. We generalize that model to the one in eq. (4.8), which allows for both deterministic and random bits.

$$\begin{aligned} c[n](z) &= c_{g_c}[n] \cdot z^{g_c} + c_{g_c-1}[n] \cdot z^{g_c-1} + \cdots && (4.7)\\ &= 1 \cdot z^{g_{det}} + 0 \cdot z^{g_{det}-1} + 0 \cdot z^{g_{det}-2} + \cdots \\ &\quad + c_{g_{ran}}[n] \cdot z^{g_{ran}} + c_{g_{ran}-1}[n] \cdot z^{g_{ran}-1} + \cdots && (4.8) \end{aligned}$$

Define:

$$g_{ran} = \max\{i | c_i \sim \text{Bernoulli}(\frac{1}{2}), g_c \geq i \geq -\infty\}. \tag{4.9}$$

g_{ran} is the index of the highest random bit of $c[n](z)$ $\forall n$. g_{ran} does not vary with n. On the other hand, the value of the bit at level g_{ran}, i.e. $c_{g_{ran}}[n]$, does. Further, $c_i[n] \sim$ Bernoulli$(\frac{1}{2}), \forall i \leq g_{ran}$ and the realizations of these bits are unknown to the transmitter and the receiver. If $g_{ran} < g_c$, then we also define g_{det} as the highest deterministic bit-level of c. $c_{g_{det}} = 1$. All bits from $c_{g_{det}-1}$ to $c_{g_{ran}+1}$ are fixed to be 0, and these bits are known to the transmitter and receiver [5].

$$g_{det} = \begin{cases} \max\{i | c_i = 1\} \text{ if } g_{ran} < g_c \\ g_{ran}, \text{ otherwise.} \end{cases} \tag{4.10}$$

By construction, we have $g_{det} \geq g_{ran}$.

This setup is illustrated in Fig. 4.2. The x_i's represent the state and y_i's are the received signal at time n. The gain $c[n]$ has one deterministic bit $c_1[n] = 1$. Bits $c_0[n], c_{-1}[n] \cdots$ are unknown Bernoulli$-(\frac{1}{2})$ random variables.

[5] The bits are fixed to zero for simplicity. This can be done without loss of generality, since the information the number of bits about $x[n]$ that can be decoded from $y[n]$ depends on the number of deterministic bits of $c[n](z)$ not the specific values. As long as $c_{g_{det}=1}$ there are enough independent equations to decode the bits.

Consider the figure on the left, at time n. Then $y_3[n] = x_2[n]$ can be observed, and is a clean bit. Bit-levels $y_2[n]$ and lower are contaminated by the multiplicative noise. The *clean* level of $x[n]$, i.e. $x_2[n]$ can be perfectly estimated, but the observer can extract no information from the *contaminated* levels. So $x_1[n]$ and lower are invisible and cannot be estimated.

The visibility of the estimator does not change at time $n+1$: it can only learn the value of the top bit. Because $x_3[n+1] = x_2[n] = 1$, the estimator learns the exact same information again. The level of contamination is increasing as the magnitude of the state (i.e. the maximum bit-level above noise) increases and hence the state estimation error grows unboundedly.

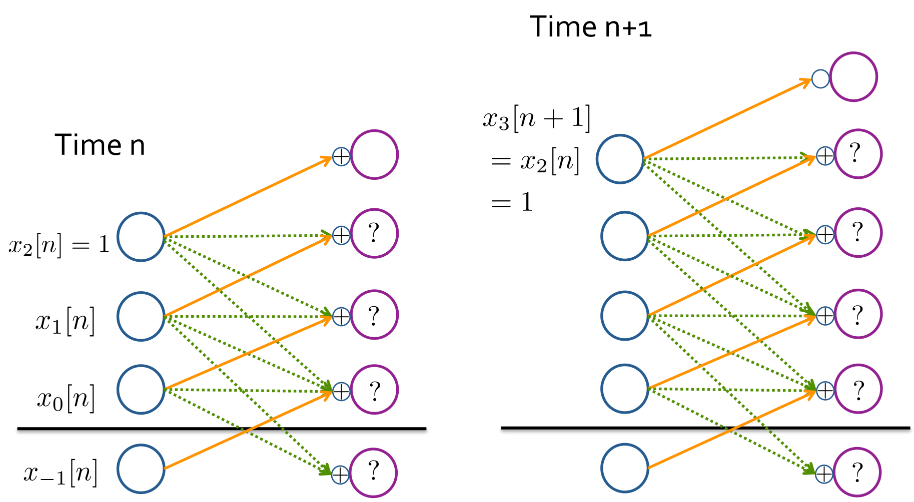


Figure 4.2: Since the stack of contaminated bits goes to infinity, so must the estimation error.

Consider on the other hand, Fig. 4.3. A controller can knock-off the top clean bit-level $x_2[n]$ at each time step, thus keeping the state bounded. As long as the rate of increase of bit-levels is not higher than the ability to control, the state can be stabilized.

Lemma 4.3.1. *$x_{d_{n+1}-i} = x_{d_n-i}$, $0 \leq i \leq d_n$, for every n.*
Proof: From eq. (4.5) we have:

$$x[n+1](z) = a \cdot x[n](z) \tag{4.11}$$
$$= z^{g_a} \cdot (x_{d_n}[n]z^{d_n} + x_{d_n-1}[n]z^{d_n-1} + \cdots + x_0[n] + x_{-1}z^{-1}\cdots) \tag{4.12}$$
$$= x_{d_n}[n]z^{d_n+g_a} + x_{d_n-1}[n]z^{d_n+g_a-1} + \cdots + x_0[n] + x_{-1}z^{-1}\cdots \tag{4.13}$$
$$= x_{d_{n+1}}[n]z^{d_{n+1}} + x_{d_{n+1}-1}[n]z^{d_n-1} + \cdots + x_0[n] + x_{-1}z^{-1}\cdots \tag{4.14}$$

We know that $d_{n+1} = d_n + g_a$. Then, comparing the coefficients of the polynomials, $x_{d_{n+1}-i} = x_{d_n-i}$. ■

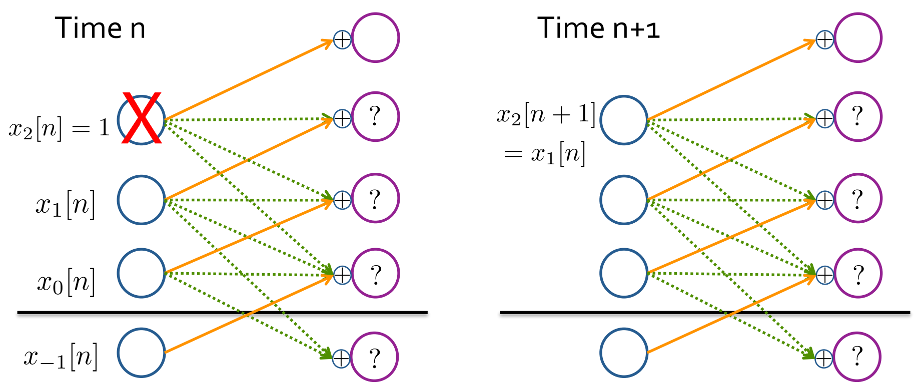


Figure 4.3: A control action can cancel out the top bit that is observable, and a fresh bit is learned every time. This strategy ensures a bounded state magnitude.

Definition 4.3.2. Let $\widehat{x}[n](z)$ be any estimate of $x[n](z)$ based on $y[n]$. Then the estimation error at time n is defined as $e[n](z) = x[n](z) - \widehat{x}[n](z)$.

Theorem 4.3.3. *Let $g_e[n]$ be the highest non-zero bit-level of $e[n](z)$. Then $\mathbb{E}[g_e[n]] \to \infty$, i.e. the degree of the error polynomial is unbounded.*

Proof: At time n, the estimator receives $y[n](z)$, and can use this to generate $\widehat{x}[n](z)$. At time 0, following Thm. 2.4.1 in Chapter 2, $I(x[n]; y[n]) = g_{det} - g_{ran} + \log d_n$. Lemma 4.3.1 implies that $x_{d_n} = x_{d_0}$. Further, $y_{d_n+g_c} = y_{d_0+g_c}$, since $y_{d_n+g_c}$ is a deterministic function of x_{d_n}. Since both d_n and x_{d_n-i} where $i < d_n$ are deterministic functions of $x[0](z)$, the estimator learns only the same bits of information over and over again, and $I(x[n]; y[n]|y[0]) = 0$. Hence the error must grow unboundedly.

■

Remark 4.3.4. The second piece of information that is in the state is the highest non-zero level of $x[n](z)$, d_n, which can be decoded from the highest non-zero level of $y[n](z)$. This is the part of the information that supposedly scales with the magnitude of the state, a la the $\log \log SNR$ result. But according to our model, d_n is a deterministic function of d_{n-1}, $d_n = d_{n-1} + g_a$, and so the value of d_{n+1} has no entropy conditioned on d_n.

Control Intuition: The carry-free model also provides an intuition for why the condition $a^2 < \frac{\mu^2+\sigma^2}{\sigma^2}$ is sufficient to control system (4.2). Taking logs of this expression, we have $\log a^2 < \log(\mu^2 + \sigma^2) - \log \sigma^2$. The carry-free model suggests that if $g_a < g_{det} - g_{ran}$ then the system can be stabilized, since at each step, we observe the top $g_{det} - g_{ran}$ bits of $x[n]$. Through these, we can extract information and then cancel them with an appropriate control. This cancellation allows new bits to be learned at the next time step. Hence, the state remains bounded. $(\mu^2 + \sigma^2)$ is the second-moment of $C[n]$ and its log is the parallel to g_{det}, and $\log \sigma^2$ captures the randomness in the system. This provides a direction to attack rate-limited control problems even when the pure estimation problem is fragile to multiplicative noise.

4.4 Estimation

To focus on the multiplicative noise bottleneck we consider a simpler system eq. (4.15) and remove the additive noise[6].

We focus on the Gaussian case with $C[n] \sim \mathcal{N}(\mu, \sigma^2)$ and $X[0] \sim \mathcal{N}(0, 1)$ since the proof is the most intuitive and elucidates the proof strategy. Sec. 4.5 remarks on extensions to other densities.

$$\begin{aligned} X[n+1] &= a \cdot X[n], \\ Y[n] &= C[n] \cdot X[n] \end{aligned} \tag{4.15}$$

This simplification reduce the problem to a tracking problem to a problem of estimating the initial state $X[0]$ using Y_0^n. However, the estimation error on $X[0]$ cannot be reduced fast enough to compensate for the rate at which the state is growing. The key idea of the proof is to bound the estimation error by a hypothesis-testing problem and use the Chernoff-Stein lower bound on the probability of error.

Theorem 4.4.1. *For the system in eq.* (4.15)*, with* $C[n] \sim \mathcal{N}(\mu, \sigma^2)$*,* $\sigma > 0$ *and* $X[0] \sim \mathcal{N}(0, 1)$*, and* $|a| > 1$ *(i.e. the system is unstable),* $\lim_{n\to\infty} \mathbb{E}[Err[n]^2] = \infty$*.*

The ideas from these results can be generalized to more general densities on $C[n]$ and $X[0]$ (see Sec. 4.5).

[6] It is tempting to think that by removing the additive noise to get system (4.15), we have reduced the problem to a known one with additive observation noise as below. Indeed, if we assume the quantities are positive and take logs:

$$\log Y[n] = \log X[n] + \log C[n].$$

One can use the average $\frac{1}{n}\sum_i (Y[i] - i \log a)$ (with appropriate scaling to account for $\log C[n]$) as an estimator for the initial state. The central limit theorem then says that the error on $\log X[0]$ decays as $\frac{1}{\sqrt{n}}$. However, this is not sufficient for us to get a tight lower-bound on the MMSE error for $X[0]$: at best the logarithmic bound only helps bound the percentage error (ratio between $X[0]$ and $\widehat{X}[0]$). For large values of the argument, the log function is flat and does not discriminate between its inputs, so a bound on $|\log X[0] - \log \widehat{X}[0]|$ cannot provide a tight bound on the error $\mathbb{E}[Err[n]]$. On the other hand, when $X[0]$ is close to 0, the steep slope of log can allow for very accurate estimation of $X[0]$. This second situation is what prevents us from appropriating the log analysis for the lower-bound on the error. Similarly, results from [120] also cannot directly apply.

Further, the central limit theorem can only bound the one particular estimator at hand, and does not help us bound the other non-linear estimators that might perform better.

4.4.1 Lemmas

Before we go into the proof of Thm. 4.4.1 we state a series of lemmas. Proofs of Lemmas 4.4.2, 4.4.3, 4.4.4 are discussed in the Appendix.

Lower bounds for estimation problems have been widely studied in the statistics literature, for instance in [111, 50] among many other references. However, many of these standard techniques to lowerbound the Bayesian MMSE error do not work in this particular case as is explored in Sec. 4.4.4. The key step is to reduce the estimation problem to a binary hypothesis testing problem. With this, we can use a modification of Stein's Lemma [25, p. 383] to provide the necessary lowerbound for the Bayesian problem.

Le Cam's lemma provides another lowerbound for the minimax error of an estimator. A modification of this bound for the Bayesian case as presented by Guntuboyina in [50] can be used to provide an alternative proof as well (see Sec. 4.4.3).

Lemma 4.4.2 states that we can lower-bound the MMSE of a Gaussian random variable by the MMSE of a binary random variable. While there are some details, the crux of the argument says that a genie giving the estimator more information about the parameter Θ can only help the scaling of the MMSE error. This is the first step to reduce the problem to a hypothesis-testing problem.

Lemma 4.4.2. *Let $Y[0], Y[1], \cdots Y[n] \sim \mathcal{P}_\Theta$. Θ is the parameter for the distribution and $\Theta \sim \mathcal{N}(0,1)$. Let $\mathcal{F}_n$ be the σ-algebra defined by $Y[0], Y[1], \cdots Y[n]$. Given fixed r_0, r_1, ϵ such that $r_1 > r_0$ and $r_1 - r_0 > \epsilon > 0$, there exist random variables R, R_1, C and a σ-algebra $\mathcal{H}_n$, such that*

- *$R_1 \sim$ Uniform$[r_0, r_1 - \epsilon]$*
- *R is a discrete random variable conditioned on R_1. $R = R_1$ with probability $\frac{1}{2}$ and $R = R_1 + \epsilon$ with probability $\frac{1}{2}$.*
- *C is a Bernoulli-(γ) random variable, where $0 < \gamma < 1$ is a constant that depends on r_0, r_1 but does not depend on ϵ.*
- *$\mathcal{H}_n = \mathcal{F}_n \cup \mathcal{F}_{R_1} \cup \mathcal{F}_C$, where $\mathcal{F}_{R_1}, \mathcal{F}_C$ are the σ-algebras generated by R_1, C respectively.*
- *$\min_{\widehat{\Theta} \in \mathcal{F}_n} \mathbb{E}|\Theta - \widehat{\Theta}|^2 \geq \min_{\widehat{R} \in \mathcal{H}_n} \gamma \cdot \mathbb{E}|R - \widehat{R}|^2$.*

The second Lemma 4.4.3 shows that the MMSE for a binary random variable can increase by at most a factor of four if the estimator is restricted to the two values in the range of the random variable.

Lemma 4.4.3. *Let S be a discrete random variable, with distribution $\mathcal{P}_S(\cdot)$ that takes values only on two points $s_1 < s_2$. $\mathcal{P}_S(s_1) > 0$, $\mathcal{P}_S(s_2) > 0$ and $\mathcal{P}_S(s) = 0\ \forall s \neq s_1, s_2$. Let $\mathcal{H}_n$ be* *the σ-algebra of the observations. Then,*

$$\min_{\widehat{S}\in\mathcal{H}_n}\mathbb{E}[|S-\widehat{S}|^2] \geq \frac{1}{4}\min_{\widehat{S}\in\mathcal{H}_n,\widehat{S}\in\{s_1,s_2\}}\mathbb{E}[|S-\widehat{S}|^2]. \tag{4.16}$$

Lemma 4.4.4 uses the Chernoff-Stein Lemma to calculate the probability of type-2 error in a hypothesis-testing problem when the type-1 error is small and the two hypotheses are getting closer and closer together. If the KL-divergence between the hypotheses is small enough, the probability of error is constant.

Lemma 4.4.4. *Let S be a random variable taking values s_1 and $s_2 \neq 0$ with equal probability, $\mathcal{P}_S(s_1) = \mathcal{P}_S(s_2) = \frac{1}{2}$, and let $Y[1], Y[2], \cdots Y[n] \sim \mathcal{N}(\mu S, \sigma^2 S^2) = \mathcal{Q}$. $\sigma \neq 0$. Let $\mathcal{P}_1$ be $\mathcal{N}(\mu s_1, \sigma^2 s_1^2)$ and $\mathcal{P}_2$ be $\mathcal{N}(\mu s_2, \sigma^2 s_2^2)$. Consider the hypothesis test between $H_1 : \mathcal{Q} = \mathcal{P}_1$ and $H_2 : \mathcal{Q} = \mathcal{P}_2$. H_1, H_2 have priors $P(H_1) = P(H_2) = \frac{1}{2}$ and $g(Y_0^n)$ is the decision rule used. Then define the probabilities of error as $\alpha_n = P(g(Y_0^n) = s_2|H_1 \text{ true})$ and $\beta_n = P(g(Y_0^n) = s_1|H_2 \text{ true})$.*

If $s_2 - s_1 = \frac{1}{n}$, $s_1 < s_2$, and $0 \leq \alpha_n < \frac{1}{n^{1+\zeta}} = \delta_n$, for some $\zeta > 0$, then, $\lim_{n\to\infty}\beta_n \geq e^{-\frac{1}{s_2}}2^{\frac{1}{s_2}} = \kappa > 0$, a constant that does not depend on n.

Remark 4.4.5. The central limit theorem states that the error of the averaging estimator decays as $\frac{1}{\sqrt{n}}$, and it is natural to think that $(s_2 - s_1)$ scaling as $\frac{1}{\sqrt{n}}$ would be the right choice to get a constant probability of error. However, this turns out to be a little too slow, and in this case the probability of error can only be lower bounded by 0 using the techniques from this chapter.

4.4.2 Proof of Thm. 4.4.1

Theorem 4.4.1. *For the system in eq. (4.15), with $C[n] \sim \mathcal{N}(\mu, \sigma^2)$, $\sigma > 0$ and $X[0] \sim \mathcal{N}(0, 1)$, and $|a| > 1$ (i.e. the system is unstable), $\lim_{n\to\infty}\mathbb{E}[Err[n]^2] = \infty$.*

Proof: Recall $X[0] \sim \mathcal{N}(0, 1)$, $C[n] \sim \mathcal{N}(\mu, \sigma^2)$ i.i.d..

Let $\mathcal{F}_n$ be the σ-algebra defined by $Y[0], Y[1], \cdots Y[n]$.

$$\begin{aligned}
&\min_{\widehat{X}[n+1]\in\mathcal{F}_n}\mathbb{E}|X[n+1]-\widehat{X}[n+1]|^2 \\
=&\min_{\widehat{X}[n+1]\in\mathcal{F}_n}\mathbb{E}|a^{n+1}X[0]-\widehat{X}[n+1]|^2 && (4.17)\\
=&a^{2(n+1)}\min_{\widehat{X}_n[0]\in\mathcal{F}_n}\mathbb{E}|X[0]-\widehat{X}_n[0]|^2. && (4.18)
\end{aligned}$$

where $\widehat{X}[n+1] = a^{n+1}\widehat{X}_n[0]$.

The only uncertainty in the system (4.15) comes from the initial state $X[0]$. For system (4.15), $X[n] = a^{n+1}X[0]$ and estimating of $X[n]$ is equivalent to estimating $X[0]$. Scaled appropriately, Y_0^n are observations drawn i.i.d. from $\mathcal{P}_{X[0]}(\cdot)$, where $X[0]$ is a parameter.

Note that $Y[n]\big|X[n] \sim \mathcal{N}(\mu X[n], \sigma^2 X[n]^2)$. Also, $\frac{Y[n]}{a^n}\big|X[0] \sim \mathcal{N}(\mu X[0], \sigma^2 X[0]^2) = \mathcal{P}_{X[0]}(\cdot)$. $\mathcal{F}_n$ is the σ-algebra generated by Y_0^n.

We start with the term $\min\limits_{\widehat{X}_n[0]\in\mathcal{F}_n} \mathbb{E}|X[0]-\widehat{X}_n[0]|^2$ and use Lemma 4.4.2 to lower-bound the estimation error of $X[0]$ by the estimation error that would result if $X[0]$ were a binary random variable.

Choose $r_0 = 1, r_1 = 2$ and $\epsilon = \frac{1}{n}$, $n \geq 1$. Lemma 4.4.2 implies that $\exists\ R, R_1, C$ and $0 < \gamma < 1$ such that R is a discrete random variable, with distribution $\mathcal{P}_R(\cdot)$, that takes only two values r and $r+\epsilon$ between r_0 and r_1, and $\mathcal{P}_R(r) = \frac{1}{2}, \mathcal{P}_R(r+\epsilon) = \frac{1}{2}$ and C is Bernoulli-(γ). This gives (4.19). $\mathcal{H}_n = \mathcal{F}_n \cup \mathcal{F}_{R_1} \cup \mathcal{F}_C$, where $\mathcal{F}_{R_1}, \mathcal{F}_C$ are the σ-algebras generated by R_1 and C respectively.

$$\begin{aligned}
&\min_{\widehat{X}_n[0]\in\mathcal{F}_n} \mathbb{E}|X[0]-\widehat{X}_n[0]|^2 \\
&\geq \min_{\widehat{R}\in\mathcal{H}_n} \gamma\cdot\mathbb{E}|R-\widehat{R}|^2 &(4.19)\\
&\geq \min_{\widehat{R}\in\mathcal{H}_n, \widehat{R}\in\{r,r+\epsilon\}} \frac{\gamma}{4}\cdot\mathbb{E}|R-\widehat{R}|^2 &(4.20)\\
&= \frac{\gamma}{4}\cdot p_e(n,\epsilon)\cdot\epsilon^2 &(4.21)\\
&\geq \frac{\gamma}{4}\cdot\frac{1}{2}\cdot(\alpha_n+\beta_n)\frac{1}{n^2}. &(4.22)
\end{aligned}$$

Eq. (4.20) states that the MMSE error considered can only be increased by a factor of 4 by reducing the domain of the estimator $\widehat{R}$ to $\{r, r+\epsilon\}$ and follows from Lemma 4.4.3. To see eq. (4.21), note that the probability of error $p_e(n,\epsilon) = P(\widehat{R} = r, R = r+\epsilon) + P(\widehat{R} = r+\epsilon, R = r)$, and hence, $\mathbb{E}|R-\widehat{R}|^2 = p_e(n,\epsilon)\cdot\epsilon^2$ if $\widehat{R}\in\{r, r+\epsilon\}$.

For eq. (4.22) define $\alpha_n = P(\widehat{R} = r|R = r+\epsilon)$ and $\beta_n = P(\widehat{R} = r+\epsilon|R = r)$. Then, $p_e(n,\epsilon) = \frac{1}{2}(\alpha_n+\beta_n)$.

At this point, the estimation problem is effectively reduced to a binary-decision problem between the two hypotheses $R = r$ and $R = r+\epsilon$. The mean-squared error in both cases is symmetric and equal to ϵ^2. $\widehat{R}$ is an estimator restricted to two values. The Chernoff-Stein lemma characterizes the probability of error of such a binary-decision problem. This lemma is used to prove Lemma 4.4.4 which gives us the key step below. Let $\delta_n = \frac{1}{n^{1+\zeta}}$, $\zeta > 0$. First, consider the case where $\alpha_n > \delta_n$,

$$\begin{aligned}
&\lim_{n\to\infty} a^{2(n+1)} \min_{\widehat{X}_n[0]\in\mathcal{F}_n} \mathbb{E}|X[0]-\widehat{X}_n[0]|^2 \\
&\geq \lim_{n\to\infty} a^{2(n+1)}\cdot\frac{\gamma}{4}\cdot\frac{1}{2}\cdot(\alpha_n+\beta_n)\cdot\frac{1}{n^2} &(4.23)\\
&\geq \lim_{n\to\infty} a^{2(n+1)}\cdot\frac{\gamma}{4}\cdot\frac{1}{2}\cdot\frac{1}{n^{1+\zeta}}\cdot\frac{1}{n^2} &(4.24)\\
&= \infty. &(4.25)
\end{aligned}$$

On the other hand if $\alpha_n < \delta_n$, by Lemma 4.4.4 we have $\lim_{n\to\infty} \beta_n > \kappa$

$$\begin{aligned}
&\lim_{n\to\infty} a^{2(n+1)} \min_{\widehat{X}_n[0]\in\mathcal{F}_n} \mathbb{E}|X[0] - \widehat{X}_n[0]|^2 \\
&\geq \lim_{n\to\infty} a^{2(n+1)} \cdot \frac{\gamma}{4} \cdot \frac{1}{2} \cdot \beta_n \cdot \frac{1}{n^2} && (4.26)\\
&> \kappa \lim_{n\to\infty} a^{2(n+1)} \cdot \frac{\gamma}{4} \cdot \frac{1}{2} \cdot \frac{1}{n^2} && (4.27)\\
&= \infty. && (4.28)
\end{aligned}$$

Since $\kappa > 0$ is a constant the exponential term dominates as $n \to \infty$ which gives eq. (4.28). ■

Remark 4.4.6. This proof can be extended to include the case with a driving disturbance $W[n]$ and an additive noise term $V[n]$ as in eq. (4.1) using a genie that just gives the disturbance and the noise as side information to the estimator. The carry-free models imply that the result should be true in the additive noise case as well.

4.4.3 Alternate proof using Le Cam's lemma

We can use the technique of reducing the problem to a hypothesis testing problem and apply a version of Le Cam's lemma to get the result as well.

Let Θ denote a parameter space. Let $L(\theta_1, \theta_2)$ denote a non-negative loss function on this space. X is an observation whose distribution $\mathcal{P}_\theta$ depends on the realization θ of unknown parameter Θ. $\mathcal{P}_\theta$ is a probability measure on a space $\mathcal{X}$.

Definition 4.4.7. Minimax error: The minimax error R_{minimax} for a problem is defined by

$$R_{\text{minimax}} = \inf_{\mathfrak{d}} \sup_{\theta\in\Theta} \mathbb{E}[L(\theta, \mathfrak{d}(X))\big|\theta] \tag{4.29}$$

where $\mathfrak{d}$ is a function that maps from $\mathcal{X}$ to Θ.

Let $(\mathcal{X}, \mathcal{F})$ be a measurable space and let $\mathcal{P}$ and $\mathcal{Q}$ be two probability measures on $(\mathcal{X}, \mathcal{F})$.

Definition 4.4.8. Total variation distance: The total variation distance between $\mathcal{P}$ and $\mathcal{Q}$ is defined as

$$V(\mathcal{P}, \mathcal{Q}) = \sup_{A\in\mathcal{F}} |\mathcal{P}(A) - \mathcal{Q}(A)| \tag{4.30}$$

Lemma 4.4.9. *(Le Cam, 1973) Let w_1 and w_2 be two probability measures that are supported on subsets Θ_1 and Θ_2 respectively. Also, let m_1 and m_2 denote the marginal densities with respect to w_1 and w_2 respectively, i.e. $m_i(x) = \int_{\Theta_i} \mathcal{P}_\theta(x) w_i(d\theta)$ for $i = 1, 2$. Then, the minimax error in the estimation of parameter θ, $R_{minimax}$, satisfies [50]:*

$$R_{minimax} \geq \frac{1}{2} d(\Theta_1, \Theta_2) V(m_1, m_2). \tag{4.31}$$

Here, $d(\Theta_1, \Theta_2) = \inf\{d(\theta_1, \theta_2) : \theta_1 \in \Theta_1, \theta_2 \in \Theta_2\}$, and $V(\cdot, \cdot)$ is the total-variation distance.

Theorem 4.4.1. *For the system in eq. (4.15), with $C[n] \sim \mathcal{N}(\mu, \sigma^2)$, $\sigma > 0$ and $X[0] \sim \mathcal{N}(0, 1)$, and $|a| > 1$ (i.e. the system is unstable), $\lim_{n\to\infty} \mathbb{E}[Err[n]^2] = \infty$.*

Proof: (Thm. 4.4.1) The steps of this proof are identical to the proof in Sec. 4.4.2 until eq. (4.20). Once the problem has been reduced to a hypothesis-testing problem for R we can use Le Cam's lemma.

Define $\mathcal{P}_R(r) = \mathcal{P}(R)\big|_{R=r}$ and $\mathcal{P}_R(r+\epsilon) = \mathcal{P}(R)\big|_{R=r+\epsilon}$. Further, define $\mathcal{P}_{Y|R}(y|r) = \mathcal{P}(y|R)\big|_{R=r}$ and $\mathcal{P}_{Y|R}(y|r+\epsilon) = \mathcal{P}(y|R)\big|_{R=r+\epsilon}$. Let $d(r_1, r_2) = \inf\limits_{a\in\mathbb{R}}\{(r_1-a)^2 + (r_2-a)^2\}$. Then following the proof of Le Cam's lemma in Chapter 2 of Guntuboyina [50]:

$$\min_{\widehat{R}\in\mathcal{H}_n, \widehat{R}\in\{r,r+\epsilon\}} \left(\mathbb{E}|r-\widehat{R}|^2 + \mathbb{E}|r+\epsilon-\widehat{R}|^2\right) \tag{4.32}$$

$$= \min_{\widehat{R}\in\mathcal{H}_n, \widehat{R}\in\{r,r+\epsilon\}} \int (r-\widehat{R}(\vec{y}))^2 \mathcal{P}_{Y|R}(\vec{y}|r) d\vec{y} + \int (r+\epsilon-\widehat{R}(\vec{y}))^2 \mathcal{P}_{Y|R}(\vec{y}|r+\epsilon) d\vec{y} \tag{4.33}$$

$$\geq \min_{\widehat{R}\in\mathcal{H}_n, \widehat{R}\in\{r,r+\epsilon\}} \int \left[(r-\widehat{R}(\vec{y}))^2 + (r+\epsilon-\widehat{R}(\vec{y}))^2\right] \min\left(\mathcal{P}_{Y|R}(\vec{y}|r), \mathcal{P}_{Y|R}(\vec{y}|r+\epsilon)\right) d\vec{y} \tag{4.34}$$

$$\geq \min_{\widehat{R}\in\mathcal{H}_n, \widehat{R}\in\{r,r+\epsilon\}} \int d(r, r+\epsilon) \min\left(\mathcal{P}_{Y|R}(\vec{y}|r), \mathcal{P}_{Y|R}(\vec{y}|r+\epsilon)\right) d\vec{y} \tag{4.35}$$

$$= \min_{\widehat{R}\in\mathcal{H}_n, \widehat{R}\in\{r,r+\epsilon\}} d(r, r+\epsilon) \int \min\left(\mathcal{P}_{Y|R}(\vec{y}|r), \mathcal{P}_{Y|R}(\vec{y}|r+\epsilon)\right) d\vec{y} \tag{4.36}$$

$$= d(r, r+\epsilon) \int \min\left(\mathcal{P}_{Y|R}(\vec{y}|r), \mathcal{P}_{Y|R}(\vec{y}|r+\epsilon)\right) d\vec{y} \tag{4.37}$$

$$= d(r, r+\epsilon) \cdot \left(1 - V(\mathcal{P}_{Y|R}(\vec{y}|r), \mathcal{P}_{Y|R}(\vec{y}|r+\epsilon))\right) \tag{4.38}$$

$$\geq d(r, r+\epsilon) \cdot \left(1 - \sqrt{\frac{D(\mathcal{P}_{Y|R}(\vec{y}|r)||\mathcal{P}_{Y|R}(\vec{y}|r+\epsilon))}{2}}\right) \tag{4.39}$$

(4.38) follows from Scheffe's theorem[7] [111, p. 84, Lemma 2.1]. (4.39) follows from Pinsker's Inequality[8] [111, p. 88, Lemma 2.5].

[7]Scheffe's theorem: $V(P,Q) = 1 - \int \min(dP, dQ)$

[8]Pinsker's Inequality: $V(P,Q) \leq \sqrt{\frac{D(P||Q)}{2}}$

Now, choose $\epsilon = \frac{1}{n}$. Then, we have $d(r, r + \frac{1}{n}) = \frac{1}{4}\frac{1}{n^2}$. Further, we have that the KL-divergence between two normal distributions $\mathcal{N}_0(\mu_0, \Sigma_0)$ and $\mathcal{N}_1(\mu_1, \Sigma)$ is given by:

$$D(\mathcal{N}_0||\mathcal{N}_1) = \frac{1}{2}\left(tr(\Sigma_1^{-1}\Sigma_0) + (\mu_1 - \mu_0)^T\Sigma_1^{-1}(\mu_1 - \mu_0) - n - \ln\frac{det\Sigma_0}{det\Sigma_1}\right) \tag{4.40}$$

So we can follow from eq. (4.39) with

$$\begin{aligned} D(\mathcal{P}_{Y|R}(\vec{y}|r)||\mathcal{P}_{Y|R}(\vec{y}|r+\epsilon)) &= \frac{1}{2}\left(\frac{n\sigma^2 r}{\sigma^2(r+\epsilon)} + \frac{n\mu^2\epsilon^2}{\sigma^2(r+\epsilon)^2} - n - n\ln\frac{\sigma^2 r^2}{\sigma^2(r+\epsilon)^2}\right) &(4.41)\\ &= \frac{1}{2}\left(\frac{n\sigma^2 r}{\sigma^2(r+\frac{1}{n})} + \frac{n\mu^2\frac{1}{n^2}}{\sigma^2(r+\frac{1}{n})^2} - n - n\ln\frac{\sigma^2 r^2}{\sigma^2(r+\frac{1}{n})^2}\right) &(4.42)\\ &= \frac{1}{2}\left(\frac{nr}{(r+\frac{1}{n})} + \frac{\mu^2\frac{1}{n}}{\sigma^2(r+\frac{1}{n})^2} - n - n\ln\frac{r^2}{(r+\frac{1}{n})^2}\right) &(4.43)\\ &= \frac{1}{2}\left(\frac{nr}{(r+\frac{1}{n})} + \frac{\mu^2\frac{1}{n}}{\sigma^2(r+\frac{1}{n})^2} - n - \ln\left(1 - \frac{1}{nr+1}\right)^n\right) &(4.44)\end{aligned}$$

$$\lim_{n\to\infty}\frac{1}{2}\left(\frac{nr}{(r+\frac{1}{n})} + \frac{\mu^2\frac{1}{n}}{\sigma^2(r+\frac{1}{n})^2} - n - \ln\left(1 - \frac{1}{nr+1}\right)^n\right) = \frac{1}{2}(-1)\ln(e^{\frac{-1}{r}}) = \frac{1}{2r} \tag{4.45}$$

Hence, we have the lowerbound

$$\begin{aligned} &\lim_{n\to\infty}\min_{\widehat{X}[n+1]\in\mathcal{F}_n}\mathbb{E}|X[n+1] - \widehat{X}[n+1]|^2\\ &\geq \lim_{n\to\infty} a^{n+1}\cdot\frac{\gamma}{8}\cdot\frac{1}{4n^2}\cdot\left(1 - \frac{1}{2}\left(\frac{nr}{(r+\frac{1}{n})} + \frac{\mu^2\frac{1}{n}}{\sigma^2(r+\frac{1}{n})^2} - n - \ln\left(1 - \frac{1}{nr+1}\right)^n\right)\right) &(4.46)\\ &\geq \lim_{n\to\infty} a^{n+1}\cdot\frac{\gamma}{8}\cdot\frac{1}{4n^2}\cdot\left(1 - \frac{1}{2r}\right) &(4.47)\\ &= \infty &(4.48)\end{aligned}$$

Thus, the MMSE is unbounded. ■

4.4.4 Connections to other lowerbounds

The Bayesian Cramer-Rao bound

It would be sacrilege to not mention the Cramer-Rao bound in the context of lowerbounds for estimation problems. The traditional Cramer-Rao bound [25, p. 395], states that variance of any unbiased estimator $\widehat{\theta}$ for a parameter θ is lower bounded by the reciprocal of the Fisher information, $J(\theta)$,

$$\mathbb{E}[(\theta - \widehat{\theta})^2|\theta] \geq \frac{1}{J(\theta)}. \tag{4.49}$$

Here $J(\theta) = \mathbb{E}\left[\left(\frac{\partial}{\partial\theta}\ln f(Y;\theta)\right)^2\middle|\theta\right]$, where $\{f(Y;\theta)\}, \theta \in \Theta$ is an indexed family of densities of the random variable Y with parameter θ.

This approach is extended to Bayesian settings with a prior on θ and for biased estimators with the Van-Trees Inequality, also called the Bayesian Cramer-Rao bound [43]. This inequality provides a lower-bound on the mean-squared error in the estimation of a parameter θ, with prior π over θ.

$$\int \mathbb{E}[(\theta - \widehat{\theta})^2\big|\theta]d\pi(\theta) \geq \frac{1}{\int J(\theta)d\pi(\theta) + J(\pi)} \quad \forall\widehat{\theta} \tag{4.50}$$

where $J(\cdot)$ is the Fisher Information. Here, $J(\pi) = \int \left(\frac{\partial ln\pi(\theta)}{\partial\theta}\right)^2 \pi(\theta)d\theta$. Both the Fisher Information of the observation and the Fisher Information of the prior are involved.

In the setup here, $\theta = X[0]$ and we have π as $\mathcal{N}(0,1)$. It is well known that $J(\pi) = 1$.

To calculate $J(\theta)$ with n i.i.d. observations $Y'[i]\big|\theta = \frac{Y[i]}{a^i}\big|\theta \sim \mathcal{N}(\mu\theta, \sigma^2\theta^2)$:

$$J(\theta) = n \cdot \mathbb{E}\left[\left(\frac{\partial}{\partial\theta}\ln f_{Y[1]}(y)\right)^2\middle|\theta\right] \tag{4.51}$$

$$= n \cdot \mathbb{E}\left[\left(\frac{\partial}{\partial\theta}\ln\left(\frac{1}{\sqrt{2\pi}\sigma|\theta|}e^{\frac{-(y-\mu\theta)^2}{2\sigma^2\theta^2}}\right)\right)^2\middle|\theta\right] \tag{4.52}$$

$$= n \cdot \mathbb{E}\left[\left(\frac{-sgn(\theta)}{|\theta|} + \frac{y(y-\mu\theta)}{\sigma^2\theta^3}\right)^2\middle|\theta\right] \tag{4.53}$$

$$= n \cdot \mathbb{E}\left[\left(-\frac{1}{\theta} + \frac{y(y-\mu\theta)}{\sigma^2\theta^3}\right)^2\middle|\theta\right] \tag{4.54}$$

$$= n \cdot \left[\frac{1}{\theta^2} - \frac{2}{\sigma^2\theta^4}\left(\mathbb{E}[y^2\big|\theta] - \mu\theta\mathbb{E}[y\big|\theta]\right) + \frac{1}{\sigma^4\theta^6}\left(\mathbb{E}[y^4\big|\theta] - 2\mu\theta\mathbb{E}[y^3\big|\theta] + \mu^2\theta^2\mathbb{E}[y^2\big|\theta]\right)\right] \tag{4.55}$$

$$= n \cdot \left[\frac{1}{\theta^2} - \frac{2}{\sigma^2\theta^4}\left(\mu^2\theta^2 + \sigma^2\theta^2 - \mu^2\theta^2\right) + \right. \tag{4.56}$$

$$\left. \frac{1}{\sigma^4\theta^6}\left(\mu^4\theta^4 + 6\mu^2\theta^2\sigma^2\theta^2 + 3\sigma^4\theta^4 - 2\mu\theta(\mu^3\theta^3 + 3\mu\theta\sigma^2\theta^2) + \mu^2\theta^2(\mu^2\theta^2 + \sigma^2\theta^2)\right)\right] \tag{4.57}$$

$$= n \cdot \left[\frac{1}{\theta^2} - \frac{2}{\sigma^2\theta^4}(\sigma^2\theta^2) + \frac{1}{\sigma^4\theta^6}\left(3\sigma^4\theta^4 + \mu^2\sigma^2\theta^4\right)\right] \tag{4.58}$$

$$= n \cdot \left[\frac{1}{\theta^2} - \frac{2}{\theta^2} + \frac{1}{\sigma^2\theta^2}\left(3\sigma^2 + \mu^2\right)\right] \tag{4.59}$$

$$= n \cdot \left[\frac{1}{\theta^2}\left(\frac{2\sigma^2 + \mu^2}{\sigma^2}\right)\right] \tag{4.60}$$

Now, let us consider,

$$\cdot \int_{-\infty}^{\infty} J(\theta) d\pi(\theta) = n \cdot \int_{-\infty}^{\infty} \left(\frac{1}{\theta^2} \left(\frac{2\sigma^2 + \mu^2}{\sigma^2} \right) \right) \frac{1}{\sqrt{2\pi}} e^{\frac{-\theta^2}{2}} d\theta \tag{4.61}$$

$$= \infty \tag{4.62}$$

As a result, a direct application of the Van-Trees inequality only provides a trivial lower bound of 0 for the MMSE. For this family of distributions, the value $\theta = 0$ removes all randomness from the system, and so the Fisher information is dominated by this one highly informative member of the family.

Ziv-Zakai bounds

The Ziv-Zakai family of bounds relate the mean squared error of an estimation problem to the probability of error in a binary hypothesis testing problem, as we do in the proof here. The original Ziv-Zakai bounds were derived for the estimation of a parameter with a uniform prior [124, 19]. This bound was extended to include general priors by Bell et al. [10]. However, we cannot evaluate this extended bound, called the Bell-Ziv-Zakai bound, to provide a closed form lowerbound in this case. Let $\pi(\theta)$ be the prior on the parameter θ.

The bound states that the MMSE in the estimation of a parameter Θ can be lowerbounded as:

$$\mathbb{E}[(\theta - \widehat{\theta})^2] = \frac{1}{2} \int_0^{\infty} P\left(|\theta - \widehat{\theta}| \geq \frac{h}{2} \right) h \; dh \tag{4.63}$$

$$P\left(|\theta - \widehat{\theta}| \geq \frac{h}{2} \right) \geq \int_{-\infty}^{\infty} (\pi(x) + \pi(x+h))\, p_e(x, x+h) \; dx \tag{4.64}$$

where $p_e(x, x+h)$ is the minimum probability of error obtained from the optimum likelihood ratio test to distinguish between x and $x+h$ using the observations.

The solution we provide uses eq. (4.63) in the regime where $h \to 0$. This lets us ignore the prior on x and thus we do not need to average over x to calculate the bound as in eq. (4.64). Here, $\pi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ and $p_{e(x,x+h)} = 2^{-nD(\mathcal{P}_1 || \mathcal{P}_2)}$, with

$$D(\mathcal{P}_1 || \mathcal{P}_2) = \log \frac{x+h}{x} + \frac{1}{2} \left(\frac{x^2}{(x+h)^2} \left(\frac{\mu^2 + \sigma^2}{\sigma^2} \right) - 1 \right) + \frac{\mu^2}{\sigma^2} \left(-\frac{x}{x+h} + \frac{1}{2} \right). \tag{4.65}$$

While this integral can be evaluated numerically, it does not have a clean closed form expression that we could use to obtain a bound.

4.4.5 Information bottleneck approach

Consider system (4.1). As $n \to \infty$ the power of $X[n]$ increases to infinity and it is feasible that the observations $Y[n]$ get more and more informative about $X[n]$ since the SNR increases to infinity. But the theorem below shows that the mutual information, $I(X[n]; Y[n])$, is in fact bounded, even as $n \to \infty$.

Theorem 4.4.10. *$I(X[n]; Y[n]) \leq \frac{1}{2}\log(\frac{\mu^2}{\sigma^2}+2) + 0.48$.*

The proof is given in the Appendix 4.8.9. A similar result is also shown in [21] for non-coherent communication.

This says the amount of new information an estimator can gain in any single time step is bounded. However, this approach is not sufficient to prove Thm. 4.4.1. If the estimator learns $\frac{1}{2}\log(\frac{\mu^2}{\sigma^2}+2) + 0.48$ bits each time, as time grows to infinity the estimator can accumulate an infinite number of bits about the state. However, we know from the deterministic model picture that the estimator goes not learn new bits at each time step, it keeps learning the same bits over and over again. A simple mutual information bound is unable to capture this. Partly this comes from the fact that the system must send the raw state value across the non-coherent channel. The ability to encode, and this send new bits across the channel at each time would allow the system to be tracked, but this is not possible in the setup we have.

4.5 Other distributions for $C[n]$

Is there something special about the Gaussian distribution on $C[n]$ that leads to the system being non-trackable? In particular, does the inestimability of the system stem from the fact that the estimator does not know the sign of $C[n]$? It turns out this is not the case. The main results of this chapter are true for any continuous distribution on $C[n]$ as well as general densities on $X[0]$. Below we work this out first for an exponential density and then a uniform density for $C[n]$.

4.5.1 Exponential $C[n]$

For the system in eq. (4.15) but with $C[n] \sim Exp(\lambda), \lambda > 0$, we have the following result.

Theorem 4.5.1. *For the system eq. (4.15), with $C[n] \sim Exp(\lambda), \lambda > 0$, $X[0] \sim \mathcal{N}(0,1)$, and $|a| > 1$ (i.e. the system is unstable), $\lim_{n\to\infty} \mathbb{E}[Err[n]^2] = \infty$.*

In this case we have, $\frac{Y[n]}{a^n} \sim Exp(\frac{\lambda}{x[0]})$. The rest of the proof follows exactly as in the case of Gaussian $C[n]$, except with Lemma 4.5.2 applied instead of Lemma 4.4.4. The proofs for these lemmas are given in the Appendix.

Lemma 4.5.2. *Let $\eta[1], \eta[2], \cdots \eta[n] \sim Exp(\frac{\lambda}{S}) = \mathcal{Q}$, where S is a random variable taking values s_1 and $s_2 \neq 0$ with equal probability. Let $\mathcal{P}_1 = Exp(\frac{\lambda}{s_1})$ and $\mathcal{P}_2 = Exp(\frac{\lambda}{s_2})$. Consider the hypothesis test between the hypothesis $H_1 : \mathcal{Q} = \mathcal{P}_1$ and hypothesis $H_2 : \mathcal{Q} = \mathcal{P}_2$. H_1, H_2 have priors $P(H_1) = P(H_2) = \frac{1}{2}$ and $g(\eta_1^n)$ is the decision rule used. Then define the probabilities of error as $\alpha_n = P(g(\eta_1^n) = s_2|H_1 \text{ true})$ and $\beta_n = P(g(\eta_1^n) = s_1|H_2 \text{ true})$. If $(s_2 - s_1)^2 = \frac{1}{n^2}$, $s_1 < s_2$, and $0 \leq \alpha_n < \frac{1}{n^{1+\zeta}}$, $\zeta > 0$, then, $\lim_{n\to\infty} \beta_n =\geq e^{\frac{1}{s_2}} 2^{\frac{1}{s_1 s_2}} = \kappa > 0$, where κ is a constant that does not depend on n.*

4.5.2 Uniform $C[n]$

The uniform distribution offers a challenge that the exponential and Gaussian distributions did not: a bounded support set. This bounded support set means that there is always a small probability that the observer will get a very highly informative sample which immediately distinguishes between the two hypotheses. As a result, the construction of the hypothesis testing problem in Lemma 4.5.4 the serves are a lowerbound to the estimation problem depends on the value of a. This is not required for the exponential or Gaussian distributions.

Theorem 4.5.3. *For the system eq.* (4.15)*, with* $C[n] \sim$ *Uniform*$[\mu_1, \mu_2], a \neq b$*,* $X[0] \sim f_X(x)$ *such that* $\exists [r_0, r_1], r_1 > r_0$ *such that* $\exists 0 < \gamma < 1$ *s.t* $f_X(x) > \frac{\gamma}{r_1 - r_0}$ $\forall\, x \in [r_0, r_1]$*, and* $|a| > 1$ *(i.e. the system is unstable),* $\lim_{n\to\infty} \mathbb{E}[Err[n]^2] = \infty$*.*

Note that the theorem above allows for a very general distribution on $X[0]$. All that is important is that a small uniform density can fit inside the density of a uniform to be able to apply the relevant lemmas. The proof is similar to the proof of Thm. 4.4.1, with Lemma 4.5.4 replacing Lemma 4.4.2 and Lemma 4.5.5 replacing Lemma 4.4.4. The main difference in the proof comes through the application of Lemma 4.5.5. The probability of error when distinguishing between two probability distributions with disjoint support can be 0 even with just a finite number of observations. For instance, for the uniform distributions in Fig. 4.4 a single observation in $[\mu_1 s_1, \mu_1 s_2]$ is sufficient to distinguish the underlying distributions. A key aspect of the proof is choosing hypotheses that are close enough to have large probability in the overlap region.

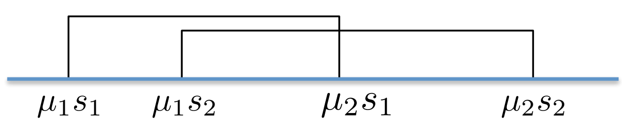


Figure 4.4: Two overlapping uniform distributions. The key to the proof is constructing a genie that maintains a sufficient amount of probability in the overlapping region $\mu_1 s_2$ to $\mu_2 s_1$.

This first lemma reduces the estimation problem to a hypothesis testing problem again. This is possible regardless of the density of $X[0]$, as long as a small uniform density can fit inside the density. Here, the two hypotheses that are generated to have a specific ratio, α, as opposed to a specific distance in the Gaussian proof.

Lemma 4.5.4. *Let* $Y[1], Y[2], \cdots Y[n] \sim \mathcal{P}_\Theta$*.* Θ *is the parameter for the distribution and let* $\Theta \sim f_\Theta(\theta)$*.* $f_\Theta(\theta)$ *is such that* $\exists\ [r_0, r_1], r_1 > r_0$ *and* $\exists\ 0 \leq \gamma \leq 1$ *such that* $f_\Theta(\theta) > \frac{\gamma}{r_1 - r_0}$ $\forall\, \theta \in [r_0, r_1]$*. Let* $\mathcal{F}_n$ *be the* σ*-algebra defined by* $Y[1], Y[2], \cdots Y[n]$*.*

For all $l > 1$, there exists α such that $l > \alpha > 1$, a random variable R, and a constant $0 < \gamma < 1$, such that

$$\min_{\widehat{\Theta} \in \mathcal{F}_n} \mathbb{E}|\Theta - \widehat{\Theta}|^2 \geq \min_{\widehat{R} \in \mathcal{F}_n} \gamma \cdot \mathbb{E}|R - \widehat{R}|^2 \tag{4.66}$$

and R is a discrete random variable that takes only two values $r, \alpha \cdot r$ such that

$$\mathcal{P}_R(r) = \frac{1}{2}, \mathcal{P}_R(\alpha \cdot r) = \frac{1}{2} \tag{4.67}$$

This second lemma characterizes the probability of error between two hypotheses s_1 and s_2 when they are observed through a multiplicative noise drawn from a uniform distribution, and is used to eventually bound the MMSE error.

Lemma 4.5.5. *Let $\mu_1, \mu_2 > 0$. Let $\eta[1], \eta[2], \cdots \eta[n] \sim U[\mu_1 S, \mu_2 S] = \mathcal{Q}$, where S is a random variable taking values s_1 and $s_2 \neq 0$ with equal probability, such that $\mu_1 s_2 < \mu_2 s_1$. Let $\mathcal{P}_1 = U[\mu_1 s_1, \mu_2 s_1]$ and $\mathcal{P}_2 = U[\mu_1 s_2, \mu_2 s_2]$. Consider the hypothesis test between the hypothesis $H_1 : \mathcal{Q} = \mathcal{P}_1$ and hypothesis $H_2 : \mathcal{Q} = \mathcal{P}_2$. H_1, H_2 have priors $P(H_1) = P(H_2) = \frac{1}{2}$ and $g(\eta_1^n)$ is the decision rule used. Then define the probabilities of error as $\alpha_n = P(g(\eta_1^n) = s_2 | H_1 \text{ true})$ and $\beta_n = P(g(\eta_1^n) = s_1 | H_2 \text{ true})$. Then, the probability of error $p_e(n) = P(g(\eta_1^n) = s_1, S = s_2) + P(g(\eta_1^n) = s_2, S = s_1) \geq \frac{1}{4} \cdot \left(\frac{\mu_2 s_1 - \mu_1 s_2}{s_1(\mu_2 - \mu_1)}\right)^n$.*

Proof: (Thm. 4.5.3) The proof follows exactly as the proof of Thm. 4.4.1. Using Lemma 4.5.4 and Lemma 4.4.3 we can get

$$\begin{aligned} &\min_{\widehat{X}[n+1] \in \mathcal{F}_n} \mathbb{E}|X[n+1] - \widehat{X}[n+1]|^2 \\ &\geq a^{2(n+1)} \cdot \frac{\gamma}{4} \cdot \min_{\widehat{R} \in \mathcal{F}_n} \mathbb{E}|R - \widehat{R}|^2 \end{aligned} \tag{4.68}$$

where R is a discrete random variable that takes only two values $r, \alpha \cdot r$ such that $\mathcal{P}_R(r) = \frac{1}{2}, \mathcal{P}_R(\alpha \cdot r) = \frac{1}{2}$, with $1 < \alpha = \frac{\mu_2}{\mu_1}\left(1 - \frac{1}{a^2}\right) + \frac{1}{a^2} - \rho$, where $0 < \rho < \frac{\mu_2}{\mu_1}\left(1 - \frac{1}{a^2}\right) + \frac{1}{a^2} - 1$. Such α and ρ exist since by Lemma 4.5.6, $1 < \frac{\mu_2}{\mu_1}\left(1 - \frac{1}{a^2}\right) + \frac{1}{a^2}$.

Lemma 4.5.6. $1 < \frac{\mu_2}{\mu_1}\left(1 - \frac{1}{a}\right) + \frac{1}{a}$.

The proof is included in the Appendix.

r and $\alpha \cdot r$ are both in the support of the original distribution $f_{X[0]}(\cdot)$ by Lemma 4.5.4.

To apply Lemma 4.5.5 with $s_1 = r$ and $s_2 = \alpha \cdot r$, we need to ensure that $\mu_1 s_2 < \mu_2 s_1$, i.e., $\alpha < \frac{\mu_2}{\mu_1}$. But since $\frac{\mu_2}{\mu_1} > 1$ we have,

$$\frac{\mu_2}{\mu_1} > \frac{\mu_2}{\mu_1} - \frac{1}{a^2}\left(\frac{\mu_2}{\mu_1} - 1\right) \tag{4.69}$$

$$> \frac{\mu_2}{\mu_1}\left(1 - \frac{1}{a^2}\right) + \frac{1}{a^2} \tag{4.70}$$

$$> \frac{\mu_2}{\mu_1}\left(1 - \frac{1}{a^2}\right) + \frac{1}{a^2} - \rho \tag{4.71}$$

$$= \alpha \tag{4.72}$$

for $a > 1$. Then Lemma 4.5.5 gives a bound on $p_e(n)$, and continuing from eq. (4.68),

$$a^{2(n+1)} \cdot \frac{\gamma}{4} \cdot \min_{\widehat{R} \in \mathcal{F}_n} \mathbb{E}|R - \widehat{R}|^2 \tag{4.73}$$

$$\geq a^{2(n+1)} \cdot \frac{\gamma}{4} \cdot p_e(n) \cdot (r(1-\alpha))^2 \tag{4.74}$$

$$\geq a^{2(n+1)} \cdot \frac{\gamma}{4} \cdot \frac{1}{4} \cdot \left(\frac{\mu_2 r - \mu_1 \alpha r}{r(\mu_2 - \mu_1)}\right)^n \cdot (r(1-\alpha))^2 \tag{4.75}$$

$$= a^{2(n+1)} \cdot \frac{\gamma}{16} \cdot \left(\frac{\mu_2 - \mu_1 \alpha}{(\mu_2 - \mu_1)}\right)^n \cdot (r(1-\alpha))^2 \to \infty \text{ as } n \to \infty \tag{4.76}$$

since $a^2 \cdot \frac{\mu_2 - \mu_1 \alpha}{(\mu_2 - \mu_1)} > 1$ by Lemma 4.5.7. ∎

Lemma 4.5.7. $a^2 \cdot \frac{\mu_2 - \mu_1 \alpha}{(\mu_2 - \mu_1)} > 1$

The proof is included in the Appendix.

Remark 4.5.8. This proof does not require that the hypothesis keep getting closer to each other as in the case of a Gaussian setup. Instead, the two hypothesis are a constant ratio apart, but this ratio depends on the growth factor a. A brute-force calculation of the probability of error allows us to bypass the use of Chernoff-Stein entirely.

4.5.3 $C[n]$ with a general continuous density

Since the uniform distribution is the prototypical continuous density, using Thm. 4.5.3, we can prove that the estimation error for a system (4.15) is unbounded with any continuous $C[n]$. The general density case reduces to to the uniform case by considering a piecewise constant approximation to the density for $C[n]$, i.e. thinking of a general density as a mixture of tiny uniform densities. Then reduce the problem to the uniform case with a genie that reveals which interval $C[n]$ was drawn from.

Theorem 4.5.9. *For the system eq. (4.15), if $C[n]$ has a continuous density with no atoms and $X[0]$ also has continuous density (possibly with atoms but such that a small uniform can fit inside the density), and $|a| > 1$ (i.e. the system is unstable), $\lim_{n\to\infty} \mathbb{E}[Err[n]^2] = \infty$.*

4.6 Control

Theorem 4.6.1. *The system in eq. (4.2), with $C[n] \sim \mathcal{N}(\mu, \sigma^2)$, $X[0] \sim \mathcal{N}(0, 1)$, and $|a| > 1$, is mean-square stabilizable if $a^2 < 1 + \frac{\mu^2}{\sigma^2}$.*

Proof: This shows that appropriate feedback control can stabilize the system in a certain parameter range in closed-loop, even though it may not be mean-square observable in open-loop. Choose $U[n] = -dY[n]$ for some $d \in \mathbb{R}$. Then

$$X[n+1] = aX[n] - dY[n] + W[n] \tag{4.77}$$
$$= (a - dC[n])\, X[n] - dV[n] + W[n]. \tag{4.78}$$

Hence,

$$\mathbb{E}[X[n+1]^2] \tag{4.79}$$
$$= \mathbb{E}\left[(a - dC[n])^2\, X[n]^2\right] + \mathbb{E}\left[d^2 V[n]^2\right] + \mathbb{E}\left[W[n]^2\right] \tag{4.80}$$
$$= (a^2 - 2ad\mu + d^2(\mu^2 + \sigma^2))\mathbb{E}[X[n]^2] + d^2 + 1. \tag{4.81}$$

If $K = a^2 - 2ad\mu + d^2(\mu^2 + \sigma^2) < 1$ the control can stabilize the system. Minimizing K gives $d = \frac{a\mu}{\mu^2+\sigma^2}$, and for this choice, $K < 1$ reduces to $a^2 < \frac{\mu^2+\sigma^2}{\sigma^2} = 1 + \frac{\mu^2}{\sigma^2}$. ■

4.7 The slow-fading case

The system in eq. (4.15) is being observed over a *fast-fading* channel, i.e. a channel that is changing faster than the time scale of the control operation. What if instead we were dealing with control in a block fading or slow fading setup, so that the observation scaling $C[n] = C$ was unknown but constant over time? Consider

$$X[n+1] = a \cdot X[n] \tag{4.82}$$
$$Y[n] = C \cdot X[n] \tag{4.83}$$

Let $X[0] \sim \mathcal{N}(0, 1)$ and $C \sim \mathcal{N}(\mu_c, \sigma_c^2)$. Then, $X[n] = a^n \cdot X[0]$, and $\frac{Y[n]}{a^n} = C \cdot X[0]$. Effectively, the observer receives the exact same observation every single time. No new bits are conveyed. We know from Thm. 4.4.10 that $I(X[0]; Y_0^n) \leq \frac{1}{2}\log(\frac{\mu^2}{\sigma^2} + 2) + 0.48$ for every

n. Since even as time grows we only learn a finite number of bits in total, the converse proof follows easily from the mutual information calculation. in this case the converse proof is easy, as $X[n] \to \infty$, there is no way to estimate $X[n]$ with finite mean squared error. When the $C[n]$s change with time (i.e. the fast fading case), we do not have identical observations. What happens though is that the extra information we learn is about the $C[n]$, which is not useful. The varying $C[n]$ case gives us the chance to average over a set of noisy of $X[0]$ and is easier in that sense. But as a result, proving the negative result converse is more difficult.

On one hand, it seems that a slow-fading system might offer an easier estimation problem, since certainly higher communication rates can be achieved in cases when the channel fading is known or can be learned. However, unlike the the communication setup, we cannot use a pilot signal here to learn the unknown "fading coefficient" in the estimation problem. The system state is randomly chosen and is what it is.

Of course, this changes entirely with the introduction of a control signal:

$$X[n+1] = a \cdot X[n] + U[n] \tag{4.84}$$
$$Y[n] = C \cdot X[n] \tag{4.85}$$

$X[0] \sim \mathcal{N}(0,1)$ and $C \sim \mathcal{N}(\mu_c, \sigma_c^2)$. For this system, the first control $U[0] = 1$ can serve as a pilot signal. Then, $Y[0] = C \cdot X[0]$ and $Y[1] = C \cdot (a \cdot X[0] + 1) = a \cdot Y[0] + C$. The controller can solve for C as $Y[1] - aY[0]$ in one time step!

In the presence of additive noise C can be learned to within a bounded error using a Kalman filter.

4.8 Appendix

4.8.1 Proof of Lemma 4.4.2

Lemma 4.4.2. *Let $Y[0], Y[1], \cdots Y[n] \sim \mathcal{P}_\Theta$. Θ is the parameter for the distribution and $\Theta \sim \mathcal{N}(0,1)$. Let $\mathcal{F}_n$ be the σ-algebra defined by $Y[0], Y[1], \cdots Y[n]$. Given fixed r_0, r_1, ϵ such that $r_1 > r_0$ and $r_1 - r_0 > \epsilon > 0$, there exist random variables R, R_1, C and a σ-algebra $\mathcal{H}_n$, such that*

- *$R_1 \sim$ Uniform$[r_0, r_1 - \epsilon]$*
- *R is a discrete random variable conditioned on R_1. $R = R_1$ with probability $\frac{1}{2}$ and $R = R_1 + \epsilon$ with probability $\frac{1}{2}$.*
- *C is a Bernoulli-(γ) random variable, where $0 < \gamma < 1$ is a constant that depends on r_0, r_1 but does not depend on ϵ.*
- *$\mathcal{H}_n = \mathcal{F}_n \cup \mathcal{F}_{R_1} \cup \mathcal{F}_C$, where $\mathcal{F}_{R_1}, \mathcal{F}_C$ are the σ-algebras generated by R_1, C respectively.*
- *$\min_{\widehat{\Theta} \in \mathcal{F}_n} \mathbb{E}|\Theta - \widehat{\Theta}|^2 \geq \min_{\widehat{R} \in \mathcal{H}_n} \gamma \cdot \mathbb{E}|R - \widehat{R}|^2$.*

Proof: Our proof shows that we can reduce the estimation problem over the entire range of $\Theta \sim \mathcal{N}(0,1)$ to an estimation problem over just two points. We do this by generating Θ as a mixture of two random variables S and T and using a genie argument.

Let $R_1 \sim \text{Uniform}[r_0, r_1-\epsilon]$. Let

$$S = \begin{cases} R_1 \text{ with prob. } \frac{1}{2} \\ R_1+\epsilon \text{ with prob. } \frac{1}{2} \end{cases} \tag{4.86}$$

for some fixed $\epsilon < r_1-\epsilon-r_0$. The density of S, $f_S(s)$, is given by:

$$f_S(s) = \begin{cases} \frac{1}{2(r_1-\epsilon-r_0)}, \ r_0 \le s < r_0+\epsilon \\ \frac{1}{(r_1-\epsilon-r_0)}, \ r_0+\epsilon \le s \le r_1-\epsilon \\ \frac{1}{2(r_1-\epsilon-r_0)}, \ r_1-\epsilon < s \le r_1 \\ 0, \ \text{otherwise.} \end{cases} \tag{4.87}$$

The density of T is essentially a Gaussian density minus $f_S(\cdot)$ with a scaling. Choose $\gamma = \frac{1}{2}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}r_1^2}$. Note γ does not depend on ϵ.

$$f_T(t) = \begin{cases} \frac{1}{1-\gamma}\left(\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}} - \gamma\frac{1}{2(r_1-\epsilon-r_0)}\right), \quad r_0 \le t < r_0+\epsilon \\ \frac{1}{1-\gamma}\left(\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}} - \gamma\frac{1}{(r_1-\epsilon-r_0)}\right), \quad r_0+\epsilon \le t \le r_1-\epsilon \\ \frac{1}{1-\gamma}\left(\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}} - \gamma\frac{1}{2(r_1-\epsilon-r_0)}\right), \ r_1-\epsilon < t \le r_1 \\ \frac{1}{1-\gamma}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}, \ \text{otherwise} \end{cases} \tag{4.88}$$

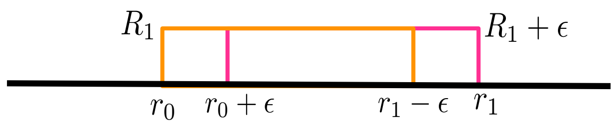
The choice of γ ensures that $f_T(t) \ge 0 \ \forall t$ and is a valid density. This mixture of S (with probability γ) and T (with probability $1-\gamma$) leads to $\Theta \sim \mathcal{N}(0,1)$. This is shown in Figures 4.5(a), 4.5(b), 4.6(a), 4.6(b).

K is a γ-biased coin which generates Θ using S, T.

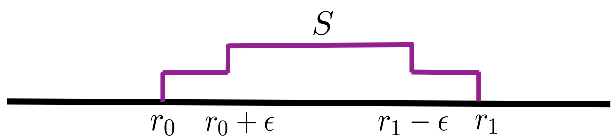
$$\Theta = \begin{cases} S \text{ if } K \text{ is heads (prob. } \gamma) \\ T \text{ if } K \text{ is tails (prob. } 1-\gamma) \end{cases} \tag{4.89}$$

Now consider the MMSE estimate of Θ, given observations Y_0^n. This error could only decrease in the event that a genie were to reveal more information about Θ. In particular, a genie observes the realization $\Theta = \theta_1$ as well as S, R_1, T, K.

If K is tails then the exact value of Θ is revealed, and $\widehat{\Theta} = \Theta$. If K is heads, then the realization of the underlying r.v. $R_1 = r$ is revealed. Let $\mathcal{F}_{R_1}$ be the σ-algebra generated by R_1. Then, $\widehat{\Theta} \in \mathcal{H}_n$, where $\mathcal{H}_n = \mathcal{F}_n \cup \mathcal{F}_{R_1}$.

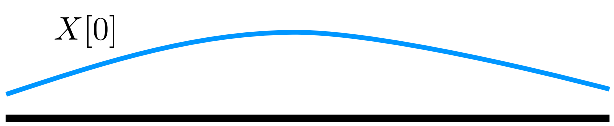


(a) Density of R_1

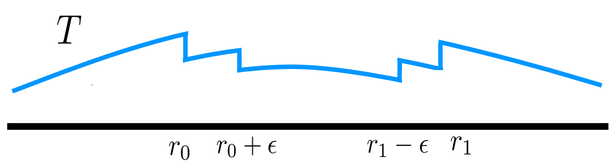


(b) Density of S

Figure 4.5: Random variables used to construct the genie side information.



(a) Caricature of the density of initial state



(b) Caricature of the density of T

Figure 4.6: Random variables used to construct the genie side information.

We are interested in the error that would be made due to confusion between r and $r+\epsilon$

after this revelation. Then,

$$\min_{\widehat{\Theta}\in\mathcal{F}_n} \mathbb{E}|\Theta - \widehat{\Theta}|^2 \tag{4.90}$$

$$\geq \min_{\widehat{\Theta}\in\mathcal{H}_n} \mathbb{E}|\Theta - \widehat{\Theta}|^2 \tag{4.91}$$

$$= \min_{\widehat{\Theta}\in\mathcal{H}_n} \gamma \cdot \mathbb{E}[|\Theta - \widehat{\Theta}|^2 | H \text{ is heads}]$$

$$+ (1-\gamma) \cdot \mathbb{E}[|\Theta - \widehat{\Theta}|^2 | H \text{ is tails}] \tag{4.92}$$

$$= \min_{\widehat{\Theta}\in\mathcal{H}_n} \gamma \cdot \mathbb{E}[|\Theta - \widehat{\Theta}|^2 | \Theta = S] + (1-\gamma) \cdot \mathbb{E}[|\Theta - \widehat{\Theta}|^2 | \Theta = T] \tag{4.93}$$

$$= \min_{\widehat{\Theta}\in\mathcal{H}_n} \gamma \cdot \mathbb{E}[|\Theta - \widehat{\Theta}|^2 | \Theta = S] \tag{4.94}$$

$$= \min_{\widehat{\Theta}\in\mathcal{H}_n} \gamma \cdot \mathbb{E}[|S - \widehat{\Theta}|^2 | \Theta = S, R_1 = r] \tag{4.95}$$

$$= \min_{\widehat{\Theta}\in\mathcal{H}_n} \frac{1}{2} \cdot \gamma \cdot \Big(\mathbb{E}[|S - \widehat{\Theta}|^2 | S = r] + \mathbb{E}[|S - \widehat{\Theta}|^2 | S = r + \epsilon]\Big) \tag{4.96}$$

(4.95) follows since when $\Theta = S$ the genie reveals $R_1 = r$. For (4.96) note that $S\big|R_1 = r$ with probability $\frac{1}{2}$ and $S\big|R_1 = r + \epsilon$ with probability $\frac{1}{2}$. So define R as:

$$R = \begin{cases} r \text{ with prob. } \frac{1}{2} \\ r + \epsilon \text{ with prob } \frac{1}{2} \end{cases} \tag{4.97}$$

If we substitute R into (4.96), we get (4.98), which gives the desired result.

$$\min_{\widehat{\Theta}\in\mathcal{F}_n} \mathbb{E}|\Theta - \widehat{\Theta}|^2$$

$$= \min_{\widehat{\Theta}\in\mathcal{H}_n} \gamma \cdot \mathbb{E}[|R - \widehat{\Theta}|^2 | \Theta = R] \tag{4.98}$$

$$= \min_{\widehat{R}\in\mathcal{H}_n} \gamma \cdot \mathbb{E}[R - \widehat{R}|^2] \tag{4.99}$$

■

4.8.2 Proof of Lemma 4.4.3

Lemma 4.4.3. *Let S be a discrete random variable, with distribution $\mathcal{P}_S(\cdot)$ that takes values only on two points $s_1 < s_2$. $\mathcal{P}_S(s_1) > 0$, $\mathcal{P}_S(s_2) > 0$ and $\mathcal{P}_S(s) = 0\ \forall s \neq s_1, s_2$. Let $\mathcal{H}_n$ be the σ-algebra of the observations. Then,*

$$\min_{\widehat{S}\in\mathcal{H}_n} \mathbb{E}[|S - \widehat{S}|^2] \geq \frac{1}{4} \min_{\widehat{S}\in\mathcal{H}_n, \widehat{S}\in\{s_1,s_2\}} \mathbb{E}[|S - \widehat{S}|^2]. \tag{4.16}$$

Proof: Let the posterior probabilities on on s_1 and s_2 by the data be q_n and $1-q_n$. Then, the MMSE estimator $\widehat{S}_n$ and associated mean squared error $Err[n] = \min_{\widehat{S}_n \in \mathcal{H}_n} \mathbb{E}[|S-\widehat{S}_n|^2]$ are

$$\widehat{S}_n = q_n s_1 + (1-q_n)s_2 \tag{4.100}$$

$$Err[n] = q_n(s_1-\widehat{S}_n))^2 + (1-q_n)(s_2-\widehat{S}_n)^2 \tag{4.101}$$

Now consider a quantization estimator $S_{quant,n} \in \{s_1, s_2\}$. $Err_{quant}[n]$ is the associated mean-squared error.

$$S_{quant,n} = \begin{cases} s_1, \text{ if } (s_1-\widehat{S}_n)^2 \leq (s_2-\widehat{S}_n)^2 \text{ i.e. } q \geq \frac{1}{2} \\ s_2, \text{ if } (s_1-\widehat{S}_n)^2 > (s_2-\widehat{S}_n)^2 \text{ i.e. } q < \frac{1}{2} \end{cases} \tag{4.102}$$

If $S_{quant,n}$ is in error, the magnitude of the error is always $(s_2-s_1)^2$. Let $p_{e,quant}$ be the probability with which $S_{quant,n}$ is in error. So

$$Err_{quant}[n] = p_{e,quant}(s_2-s_1)^2 \tag{4.103}$$

Clearly, $Err_{quant}[n] \geq Err[n]$, by the definition of $\widehat{S}_n$.

Consider the expression, $Err_{quant,genie}[n]$, which reduces the error by a factor of 4.

$$Err_{quant,genie}[n] = \frac{1}{4}Err_{quant}[n] = p_{e,quant}\left(\frac{s_1-s_2}{2}\right)^2. \tag{4.104}$$

We will show that

$$Err[n] \geq Err_{quant,genie}[n]. \tag{4.105}$$

To get (4.105), compare $Err[n]$ and $Err_{quant,genie}[n]$ on a realization by realization basis. Let $q_n > \frac{1}{2}$. Then $S_{quant,n} = s_1$. Also, $p_{e,quant} = 1-q_n$, since this is the probability with which $S = s_2$. Since $q_n > \frac{1}{2}$, we have $s_1 \leq \widehat{S}_n < \frac{s_1+s_2}{2} < s_2$. If $S = s_1$, $|Err_{quant,genie}[n]|^2 = 0$, and $|S-\widehat{S}_n|^2 = (s_1-\widehat{S}_n)^2$. So, $|Err_{quant,genie}[n]|^2 \leq |S-S_n|^2$ in this case. If $S = s_2$, $|Err_{quant,genie}[n]|^2 = \left(\frac{s_1-s_2}{2}\right)^2$ and $|S-\widehat{S}_n|^2 = (s_1-s_2)^2$. $|Err_{quant,genie}[n]|^2 \leq |S-\widehat{S}_n|^2$ in this case as well. A similar argument holds if $q \leq \frac{1}{2}$. Hence, we have $Err[n] \geq Err_{quant,genie}[n]$.

Consider now the optimal estimator for S over those estimators that only take values on $\{s_1, s_2\}$

$$S_{opt,n} = \operatorname*{argmin}_{\widehat{S} \in \mathcal{H}_n, \widehat{S} \in \{s_1,s_2\}} \mathbb{E}[|S-\widehat{S}|^2] \tag{4.106}$$

$Err_{opt}[n]$ is the associated mean squared error, and $Err_{opt,genie}[n]$ is the associated error for the optimal estimator with the help of the genie so that $Err_{opt,genie}[n] = \frac{1}{4}Err_{opt}[n]$, since the genie reduces the error by a factor of 4. $S_{opt,n}$ will have better performance than $S_{quant,n}$ and hence the probability of error $p_{e,opt} \leq p_{e,quant}$. Hence,

$$Err_{quant,genie}[n] \geq Err_{opt,genie}[n] = \frac{1}{4} Err_{opt}[n] \tag{4.107}$$

Hence, (4.105), (4.107) imply that $Err[n] \geq \frac{1}{4} Err_{opt}[n]$ i.e.

$$\min_{\widehat{S} \in \mathcal{H}_n} \mathbb{E}[|S - \widehat{S}|^2] \geq \frac{1}{4} \min_{\widehat{S} = \mathcal{H}_n, \widehat{S} \in \{s_1, s_2\}} \mathbb{E}[|S - \widehat{S}|^2] \tag{4.108}$$

■

4.8.3 Proof of Lemma 4.4.4

Lemma 4.4.4. *Let S be a random variable taking values s_1 and $s_2 \neq 0$ with equal probability, $\mathcal{P}_S(s_1) = \mathcal{P}_S(s_2) = \frac{1}{2}$, and let $Y[1], Y[2], \cdots Y[n] \sim \mathcal{N}(\mu S, \sigma^2 S^2) = \mathcal{Q}$. $\sigma \neq 0$. Let $\mathcal{P}_1$ be $\mathcal{N}(\mu s_1, \sigma^2 s_1^2)$ and $\mathcal{P}_2$ be $\mathcal{N}(\mu s_2, \sigma^2 s_2^2)$. Consider the hypothesis test between $H_1 : \mathcal{Q} = \mathcal{P}_1$ and $H_2 : \mathcal{Q} = \mathcal{P}_2$. H_1, H_2 have priors $P(H_1) = P(H_2) = \frac{1}{2}$ and $g(Y_0^n)$ is the decision rule used. Then define the probabilities of error as $\alpha_n = P(g(Y_0^n) = s_2 | H_1 \text{ true})$ and $\beta_n = P(g(Y_0^n) = s_1 | H_2 \text{ true})$.*

If $s_2 - s_1 = \frac{1}{n}$, $s_1 < s_2$, and $0 \leq \alpha_n < \frac{1}{n^{1+\zeta}} = \delta_n$, for some $\zeta > 0$, then, $\lim_{n \to \infty} \beta_n \geq e^{-\frac{1}{s_2}} 2^{\frac{1}{s_2}} = \kappa > 0$, a constant that does not depend on n.

Proof: $s_2 - s_1 = \frac{1}{n}$, thus $\frac{s_1}{s_2} = 1 - \frac{1}{n s_2}$ and $\left(\frac{s_1^2}{s_2^2} - 1 \right) = \frac{1}{n} \left(\frac{1}{n s_2^2} - \frac{2}{s_2} \right)$.

Consider any decision rule, $g(Y_0^n)$. Associated with the decision rule is an acceptance region for H_1 based on Y_0^n, i.e. the set of values of Y_0^n where $g(Y_0^n) = s_1$. Let B_n denote this acceptance region. Then $\alpha_n = \mathcal{P}_1(B_n^c) < \delta_n$, so $\mathcal{P}_1(B_n) > 1 - \delta_n$. Since $D(\mathcal{P}_1 || \mathcal{P}_2) < \infty$, the Chernoff-Stein lemma (Lem 11.8.1, Thm 11.8.3, [25], pg. 383) implies $\mathcal{P}_2(B_n) > (1 - 2\delta_n) 2^{-nD(\mathcal{P}_1 || \mathcal{P}_2) + \delta_n}$. Recall, $\delta_n = \frac{1}{n^{1+\zeta}}$.

When $\sigma \neq 0$,

$$\lim_{n \to \infty} \mathcal{P}_2(B_n) > \lim_{n \to \infty} (1 - 2\delta_n) 2^{-n(D(\mathcal{P}_1 || \mathcal{P}_2) + \delta_n)} \tag{4.109}$$

$$= \lim_{n \to \infty} (1 - \frac{2}{n^{1+\zeta}}) 2^{-\frac{n}{n^{1+\zeta}}} \lim_{n \to \infty} 2^{-nD(\mathcal{P}_1 || \mathcal{P}_2)} \tag{4.110}$$

$$= 1 \cdot 2^0 \lim_{n \to \infty} 2^{-nD(\mathcal{P}_1 || \mathcal{P}_2)} \tag{4.111}$$

$$D(\mathcal{P}_1 || \mathcal{P}_2) = \log \frac{s_2}{s_1} + \frac{1}{2} \left(\frac{s_1^2}{s_2^2} \left(\frac{\mu^2 + \sigma^2}{\sigma^2} \right) - 1 \right) + \frac{\mu^2}{\sigma^2} \left(-\frac{s_1}{s_2} + \frac{1}{2} \right)$$

$$
\begin{aligned}
&\lim_{n\to\infty} 2^{-nD(\mathcal{P}_1||\mathcal{P}_2)} \\
&= \lim_{n\to\infty} 2^{\left\{-n\left(\log\frac{s_2}{s_1}+\frac{1}{2}\left(\frac{s_1^2}{s_2^2}\left(\frac{\mu^2+\sigma^2}{\sigma^2}\right)-1\right)+\frac{\mu^2}{\sigma^2}\left(-\frac{s_1}{s_2}+\frac{1}{2}\right)\right)\right\}} && (4.112)\\
&= \lim_{n\to\infty}\left(\frac{s_1}{s_2}\right)^n\cdot 2^{\left\{-\frac{n}{2}\frac{1}{\sigma^2 s_2^2}(s_1^2(\mu^2+\sigma^2)-\sigma^2 s_2^2-2\mu^2 s_1 s_2+\mu^2 s_2^2)\right\}} && (4.113)\\
&= \lim_{n\to\infty}\left(\frac{s_1}{s_2}\right)^n\cdot 2^{\left\{-\frac{n}{2}\frac{1}{\sigma^2 s_2^2}\left(\mu^2(s_1^2+s_2^2-2s_1s_2)+\sigma^2(s_1^2-s_2^2)\right)\right\}} && (4.114)\\
&= \lim_{n\to\infty}\left(\frac{s_1}{s_2}\right)^n\cdot 2^{\left\{-\frac{n}{2}\frac{1}{\sigma^2 s_2^2}\left(\mu^2(s_1-s_2)^2-\sigma^2(s_2-s_1)(s_1+s_2)\right)\right\}} && (4.115)\\
&= \lim_{n\to\infty}\left(1-\frac{1}{ns_2}\right)^n\cdot 2^{\left\{-\frac{n}{2}\frac{1}{\sigma^2 s_2^2}\left(\mu^2\frac{1}{n^2}-\sigma^2\frac{1}{n}(s_1+s_2)\right)\right\}} && (4.116)\\
&= e^{-\frac{1}{s_2}}2^{\frac{1}{s_2}} && (4.117)
\end{aligned}
$$

Thus, $\lim_{n\to\infty}\mathcal{P}_2(B_n) > \left(\frac{2}{e}\right)^{\frac{1}{s_2}} = \kappa$.

■

4.8.4 Proof of Lemma 4.5.2

Lemma 4.5.2. *Let $\eta[1],\eta[2],\cdots\eta[n]\sim Exp(\frac{\lambda}{S})=\mathcal{Q}$, where S is a random variable taking values s_1 and $s_2\neq 0$ with equal probability. Let $\mathcal{P}_1=Exp(\frac{\lambda}{s_1})$ and $\mathcal{P}_2=Exp(\frac{\lambda}{s_2})$. Consider the hypothesis test between the hypothesis $H_1:\mathcal{Q}=\mathcal{P}_1$ and hypothesis $H_2:\mathcal{Q}=\mathcal{P}_2$. H_1,H_2 have priors $P(H_1)=P(H_2)=\frac{1}{2}$ and $g(\eta_1^n)$ is the decision rule used. Then define the probabilities of error as $\alpha_n=P(g(\eta_1^n)=s_2|H_1\ true)$ and $\beta_n=P(g(\eta_1^n)=s_1|H_2\ true)$. If $(s_2-s_1)^2=\frac{1}{n^2}$, $s_1<s_2$, and $0\le\alpha_n<\frac{1}{n^{1+\zeta}}$, $\zeta>0$, then, $\lim_{n\to\infty}\beta_n =\geq e^{\frac{1}{s_2}}2^{\frac{1}{s_1s_2}}=\kappa>0$, where κ is a constant that does not depend on n.*

Proof:

$$
\begin{aligned}
D(\mathcal{P}_1||\mathcal{P}_2) &= \int_0^\infty \frac{\lambda}{s_1}e^{-\frac{\lambda}{s_1}s}\log\frac{\frac{\lambda}{s_1}e^{-\frac{\lambda}{s_1}s}}{\frac{\lambda}{s_2}e^{-\frac{\lambda}{s_2}s}}ds && (4.118)\\
&= \int_0^\infty \frac{\lambda}{s_1}e^{-\frac{\lambda}{s_1}s}\log\frac{s_1}{s_2}ds+\int_0^\infty\frac{\lambda}{s_1}e^{-\frac{\lambda}{s_1}s}(-\lambda s)\left(\frac{1}{s_1}-\frac{1}{s_2}\right)ds && (4.119)\\
&= \log\frac{s_1}{s_2}-\frac{1}{s_2}+\frac{1}{s_1}. && (4.120)
\end{aligned}
$$

Note we have $s_2-s_1=\frac{1}{n}$ and $\frac{s_1}{s_2}=1-\frac{1}{ns_2}$.

$$\lim_{n\to\infty} 2^{-nD(\mathcal{P}_1||\mathcal{P}_2)} = \lim_{n\to\infty} 2^{-n\log\frac{s_1}{s_2}-\frac{1}{s_2}+\frac{1}{s_1}} \tag{4.121}$$

$$= \lim_{n\to\infty}\left(1-\frac{1}{ns_2}\right)^{-n}\cdot 2^{-n\frac{s_2-s_1}{s_1 s_2}} \tag{4.122}$$

$$= e^{\frac{1}{s_2}} 2^{\frac{1}{s_1 s_2}} \tag{4.123}$$

$$= \kappa \tag{4.124}$$

■

4.8.5 Proof of Lemma 4.5.4

Lemma 4.5.4. *Let $Y[1], Y[2], \cdots Y[n] \sim \mathcal{P}_\Theta$. Θ is the parameter for the distribution and let $\Theta \sim f_\Theta(\theta)$. $f_\Theta(\theta)$ is such that $\exists\ [r_0, r_1], r_1 > r_0$ and $\exists\ 0 \le \gamma \le 1$ such that $f_\Theta(\theta) > \frac{\gamma}{r_1-r_0}\ \forall\ \theta \in [r_0, r_1]$. Let $\mathcal{F}_n$ be the σ-algebra defined by $Y[1], Y[2], \cdots Y[n]$.*

For all $l > 1$, there exists α such that $l > \alpha > 1$, a random variable R, and a constant $0 < \gamma < 1$, such that

$$\min_{\widehat{\Theta}\in\mathcal{F}_n} \mathbb{E}|\Theta-\widehat{\Theta}|^2 \ge \min_{\widehat{R}\in\mathcal{F}_n} \gamma\cdot\mathbb{E}|R-\widehat{R}|^2 \tag{4.66}$$

and R is a discrete random variable that takes only two values $r, \alpha\cdot r$ such that

$$\mathcal{P}_R(r) = \frac{1}{2}, \mathcal{P}_R(\alpha\cdot r) = \frac{1}{2} \tag{4.67}$$

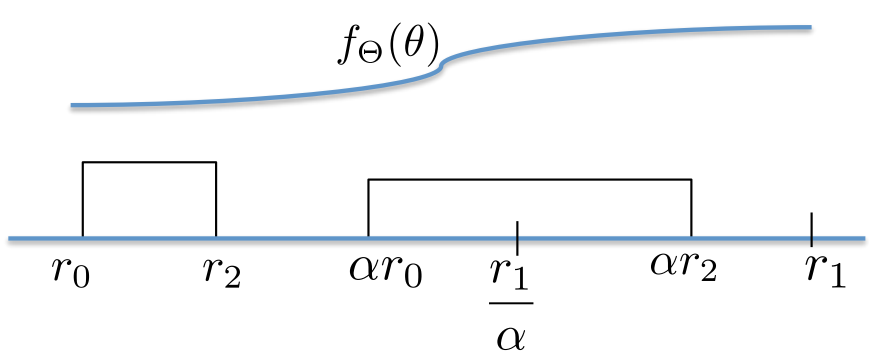


Figure 4.7: The density of S fitting inside the density of Θ.

Proof:

$\Theta \sim f_\Theta(\theta)$. Choose α and r_2 such that $r_0 < r_2 < \alpha r_0 < \frac{r_1}{\alpha} < \alpha r_2 < r_1$. This is possible when we choose an α such that $\alpha r_0 < \frac{r_1}{\alpha}$, which is always possible since α can be arbitrarily close to 1. Let $R \sim \text{Uniform}[r_0, r_2]$. Then, we flip a fair coin, K and generate the random variable S using this coin flip and R. With probability $\frac{1}{2}$, $S = R$, and with probability $\frac{1}{2}$, $S = \alpha \cdot R$ The pdf of S, $f_S(s)$, is given by:

$$f_S(s) = \begin{cases} \frac{1}{2(r_2 - r_0)}, \ r_0 \leq s < r_2 \\ \frac{1}{2\alpha(r_2 - r_0)}, \ \alpha r_0 \leq s \leq \alpha r_2 \\ 0, \ \text{otherwise} \end{cases} \tag{4.125}$$

Now we show how Θ can be generated as a mixture of S and another random variable T. The density of T is essentially the density of $X[0]$ minus the density of S. $\Theta = S$ with probability γ and $\Theta = T$ with probability $1 - \gamma$, where $0 < \gamma < 1$ and $f_\Theta(\theta) - \gamma\frac{1}{r_1 - r_0} > 0$ if $r_0 \leq \theta \leq r_1$. Note that this is possible since we have that $f_\Theta(\theta) > \frac{\gamma}{r_1 - r_0}$ when $r_0 \leq \theta \leq r_1$. To be precise the pdf of T is:

$$f_T(t) = \begin{cases} \frac{1}{1-\gamma}(f_\Theta(t) - \gamma\frac{1}{2(r_2 - r_0)}), \ r_0 \leq t \leq r_2 \\ \frac{1}{1-\gamma}(f_\Theta(t) - \gamma\frac{1}{2\alpha(r_2 - r_0)}), \ \alpha r_0 \leq t \leq \alpha r_2 \\ \frac{1}{1-\gamma}f_\Theta(t), \ \text{otherwise} \end{cases} \tag{4.126}$$

This is a valid pdf due to the choice of γ. This mixture of S and T leads to $\Theta \sim f_\Theta(\theta)$. The rest of the proof proceeds exactly as in Lemma 4.4.2.

■

4.8.6 Proof of Lemma 4.5.5

Lemma 4.5.5. *Let $\mu_1, \mu_2 > 0$. Let $\eta[1], \eta[2], \cdots \eta[n] \sim U[\mu_1 S, \mu_2 S] = \mathcal{Q}$, where S is a random variable taking values s_1 and $s_2 \neq 0$ with equal probability, such that $\mu_1 s_2 < \mu_2 s_1$. Let $\mathcal{P}_1 = U[\mu_1 s_1, \mu_2 s_1]$ and $\mathcal{P}_2 = U[\mu_1 s_2, \mu_2 s_2]$. Consider the hypothesis test between the hypothesis $H_1 : \mathcal{Q} = \mathcal{P}_1$ and hypothesis $H_2 : \mathcal{Q} = \mathcal{P}_2$. H_1, H_2 have priors $P(H_1) = P(H_2) = \frac{1}{2}$ and $g(\eta_1^n)$ is the decision rule used. Then define the probabilities of error as $\alpha_n = P(g(\eta_1^n) = s_2 | H_1$ true$)$ and $\beta_n = P(g(\eta_1^n) = s_1 | H_2$ true$)$. Then, the probability of error $p_e(n) = P(g(\eta_1^n) = s_1, S = s_2) + P(g(\eta_1^n) = s_2, S = s_1) \geq \frac{1}{4} \cdot \left(\frac{\mu_2 s_1 - \mu_1 s_2}{s_1(\mu_2 - \mu_1)}\right)^n$.*

Proof:

$$\begin{aligned} p_e(n) &= P(g(\eta_1^n) = s_1, S = s_2) + P(g(\eta_1^n) = s_2, S = s_1) &(4.127)\\ &= P(g(\eta_1^n) = s_1 | S = s_2)P(S = s_2) + P(g(\eta_1^n) = s_2 | S = s_1)P(S = s_1) &(4.128)\\ &= \frac{1}{2}(P(g(\eta_1^n) = s_1 | S = s_2) + P(g(\eta_1^n) = s_2 | S = s_1)) &(4.129) \end{aligned}$$

We expand the second term to write:

$$
\begin{aligned}
&P(g(\eta_1^n) = s_2 \big| S = s_1) && (4.130)\\
&= P(g(\eta_1^n) = s_2 \big| S = s_1, \eta_1^n \in [\mu_1 s_2, \mu_2 s_1]^n) \cdot P(\eta_1^n \in [\mu_1 s_2, \mu_2 s_1]^n \big| S = s_1) + \\
&\quad P(g(\eta_1^n) = s_2 \big| S = s_1, \eta_1^n \notin [\mu_1 s_2, \mu_2 s_1]^n) \cdot P(\eta_1^n \notin [\mu_1 s_2, \mu_2 s_1]^n \big| S = s_1) && (4.131)\\
&= P(g(\eta_1^n) = s_2 \big| S = s_1, \eta_1^n \in [\mu_1 s_2, \mu_2 s_1]^n) \cdot \left(\frac{\mu_2 s_1 - \mu_1 s_2}{s_1(\mu_2 - \mu_1)}\right)^n + \\
&\quad 0 \cdot P(\eta_1^n \notin [\mu_1 s_2, \mu_2 s_1]^n \big| S = s_1) && (4.132)\\
&= P(g(\eta_1^n) = s_2 \big| S = s_1, \eta_1^n \in [\mu_1 s_2, \mu_2 s_1]^n) \cdot \left(\frac{\mu_2 s_1 - \mu_1 s_2}{s_1(\mu_2 - \mu_1)}\right)^n && (4.133)
\end{aligned}
$$

For eq. (4.132) note $P(\eta_1^n \in [\mu_1 s_2, \mu_2 s_1]^n \big| S = s_1) = \left(\frac{\mu_2 s_1 - \mu_1 s_2}{s_1(\mu_2 - \mu_1)}\right)^n$. Also, if $\exists\ \eta[i] \notin [\mu_1 s_2, \mu_2 s_1]$, then the probability of error is 0, since even one sample outside the interval of overlap informs us about the underlying true distribution. $g(\eta_1^n) = s_1$ if $\exists\ \eta[i] \in [\mu_1 s_1, \mu_1 s_2]$, and $g(\eta_1^n) = s_2$ if $\exists\ \eta[i] \in [\mu_2 s_1, \mu_2 s_2]$. In a similar fashion we can also show:

$$
P(g(\eta_1^n) = s_1 \big| S = s_2) = P(g(\eta_1^n) = s_1 \big| S = s_2, \eta_1^n \in [\mu_1 s_1, \mu_2 s_1]^n) \cdot \left(\frac{\mu_2 s_1 - \mu_1 s_2}{s_2(\mu_2 - \mu_1)}\right)^n \quad (4.134)
$$

We would now like to understand the probability of error if all observations $\eta_1^n \in [\mu_1 s_2, \mu_2 s_1]^n$.

$\eta[i] \sim \mathcal{Q}$ for $1 \le i \le n$. We are interested in the distribution of $\eta[i]$ conditioned on the event $A = \{\eta[i] \in [\mu_1 s_2, \mu_2 s_1]\}$.

If $\mathcal{Q} = \mathcal{P}_1 = U[\mu_1 s_1, \mu_2 s_1]$, then $\eta[i] \sim \mathcal{P}_1 \big| A \sim U[\mu_1 s_2, \mu_2 s_1]$. If $\mathcal{Q} = \mathcal{P}_2 = U[\mu_1 s_2, \mu_2 s_2]$, then $\eta[i] \sim \mathcal{P}_2 \big| A \sim U[\mu_1 s_2, \mu_2 s_1]$. Thus, $\mathcal{P}_{1A} = \mathcal{P}_1 \big| A$ and $\mathcal{P}_{2A} = \mathcal{P}_2 \big| A$ are the same distribution, and are indistinguishable using the observations. By symmetry, at least one of $P(g(\eta_1^n) = s_1 \big| S = s_2, \eta_1^n \in [\mu_1 s_1, \mu_2 s_1]^n)$ and $P(g(\eta_1^n) = s_2 \big| S = s_1, \eta_1^n \in [\mu_1 s_1, \mu_2 s_1]^n)$ must be $\ge \frac{1}{2}$. Thus,

$$
\begin{aligned}
p_e(n) &\ge \frac{1}{2}\left(\frac{\mu_2 s_1 - \mu_1 s_2}{s_1(\mu_2 - \mu_1)}\right)^n \left(P(g(\eta_1^n) = s_2 \big| S = s_1, \eta_1^n \in [\mu_1 s_2, \mu_2 s_1]^n)\right) + \\
&\quad \frac{1}{2}\left(\frac{\mu_2 s_1 - \mu_1 s_2}{s_2(\mu_2 - \mu_1)}\right)^n \left(P(g(\eta_1^n) = s_1 \big| S = s_1, \eta_1^n \in [\mu_1 s_1, \mu_2 s_1]^n)\right) && (4.135)\\
&\ge \frac{1}{2} \cdot \left(\frac{\mu_2 s_1 - \mu_1 s_2}{s_1(\mu_2 - \mu_1)}\right)^n \cdot \big(P(g(\eta_1^n) = s_2 \big| S = s_1, \eta_1^n \in [\mu_1 s_2, \mu_2 s_1]^n) + \\
&\quad P(g(\eta_1^n) = s_1 \big| S = s_2, \eta_1^n \in [\mu_1 s_2, \mu_2 s_1]^n)\big) && (4.136)\\
&\ge \frac{1}{2} \cdot \left(\frac{\mu_2 s_1 - \mu_1 s_2}{s_1(\mu_2 - \mu_1)}\right)^n \cdot \left(\frac{1}{2}\right) && (4.137)\\
&= \frac{1}{4} \cdot \left(\frac{\mu_2 s_1 - \mu_1 s_2}{s_1(\mu_2 - \mu_1)}\right)^n && (4.138)
\end{aligned}
$$

■

4.8.7 Proof of Lemma 4.5.6

Lemma 4.5.6. $1 < \frac{\mu_2}{\mu_1}\left(1 - \frac{1}{a}\right) + \frac{1}{a}$.

Proof:

$$\frac{\mu_2}{\mu_1}\left(1-\frac{1}{a^2}\right)+\frac{1}{a^2} > 1 \tag{4.139}$$
$$\iff \mu_2\left(1-\frac{1}{a^2}\right)+\frac{\mu_1}{a^2} > \mu_1 \tag{4.140}$$
$$\iff \mu_2\left(1-\frac{1}{a^2}\right) > \mu_1\left(1-\frac{1}{a^2}\right) \tag{4.141}$$
$$\iff \mu_2 > \mu_1 \tag{4.142}$$

which is trivially true for $a^2 > 1$ and $\mu_1, \mu_2 > 0$. ■

4.8.8 Proof of Lemma 4.5.7

Lemma 4.5.7. $a^2 \cdot \frac{\mu_2 - \mu_1\alpha}{(\mu_2-\mu_1)} > 1$

Proof: Note that $a^2 > 0$. Then,

$$a^2 \cdot \frac{\mu_2-\mu_1\alpha}{(\mu_2-\mu_1)} > 1 \tag{4.143}$$
$$\iff \frac{\mu_2}{\mu_2-\mu_1} - \alpha\frac{\mu_1}{\mu_2-\mu_1} > \frac{1}{a^2} \tag{4.144}$$
$$\iff \alpha < \frac{1}{a^2}\frac{\mu_2-\mu_1}{-\mu_1} + \frac{\mu_2}{\mu_1} = \frac{\mu_2}{\mu_1}\left(1-\frac{1}{a^2}\right)+\frac{1}{a^2} \tag{4.145}$$

■

4.8.9 Proof of Thm. 4.4.10

Theorem 4.4.10. $I(X[n];Y[n]) \leq \frac{1}{2}\log(\frac{\mu^2}{\sigma^2}+2)+0.48$.

Proof: $X[n] \sim \mathcal{N}(0, \frac{a^{2(n+1)}-1}{a^2-1}) = \xi_n\mathcal{N}(0,1)$. Let $X[n] = \xi_n\tilde{X}[n]$, and $\tilde{X}[n] \sim \mathcal{N}(0,1)$. Hence, $Y[n] = C[n]\xi_n\tilde{X}[n] + V[n]$.

$$
\begin{aligned}
I\left(X[n];Y[n]\right) &= I\left(\xi_n\tilde{X}[n];Y[n]\right) && (4.146)\\
&= I\left(\tilde{X}[n];Y[n]\right) && (4.147)\\
&= I\left(\tilde{X}[n];\frac{Y[n]}{\xi_n\sigma}\right) && (4.148)\\
&= h\left(\frac{Y[n]}{\xi_n\sigma}\right) - h\left(\frac{Y[n]}{\xi_n\sigma}|\tilde{X}[n]\right) && (4.149)
\end{aligned}
$$

(4.147), (4.148) follow since scaling by a constant does not change mutual information. $var(\frac{Y[n]}{\xi_n\sigma}) = var(\frac{C[n]}{\sigma}\tilde{X}[n] + \frac{V[n]}{\xi_n\sigma}) = var(\frac{C[n]}{\sigma}\tilde{X}[n]) + var(\frac{V[n]}{\xi_n\sigma}) = \frac{\mu^2\cdot 1+0\cdot\sigma^2+\sigma^2\cdot 1}{\sigma^2} + \frac{1}{\xi_n^2\sigma^2}$. Hence,

$$
h\left(\frac{Y[n]}{\xi_n\sigma}\right) \leq \frac{1}{2}\log 2\pi e\left(\frac{\mu^2}{\sigma^2} + 1 + \frac{1}{\xi_n^2\sigma^2}\right). \tag{4.150}
$$

$$h\left(\frac{Y[n]}{\xi_n\sigma}|\tilde{X}[n]\right) \tag{4.151}$$
$$= \mathbb{E}\left[h\left(\frac{Y[n]}{\xi_n\sigma}\right)\middle|\tilde{X}[n]=\tilde{x}_0\right] \tag{4.152}$$
$$= \mathbb{E}\left[h\left(\frac{C[n]}{\sigma}\tilde{X}_0+\frac{V[n]}{\xi_n\sigma}\right)|\tilde{X}[n]=\tilde{x}_0\right] \tag{4.153}$$
$$= \mathbb{E}\left[h(\mathcal{N}(\frac{\mu}{\sigma}\cdot\tilde{X}_0,\tilde{X}_0^2+\frac{1}{\xi_n^2\sigma^2})|\tilde{X}[n]=\tilde{x}_0\right] \tag{4.154}$$
$$= \frac{1}{2}\log 2\pi e+\frac{1}{2}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{\frac{-x^2}{2}}\left(\log(x^2+\frac{1}{\xi_n^2\sigma^2})\right)dx \tag{4.155}$$
$$= \frac{1}{2}\log 2\pi e+\int_{0}^{1}\frac{1}{\sqrt{2\pi}}e^{\frac{-x^2}{2}}\left(\log(x^2+\frac{1}{\xi_n^2\sigma^2})\right)dx+\int_{1}^{\infty}\frac{1}{\sqrt{2\pi}}e^{\frac{-x^2}{2}}\left(\log(x^2+\frac{1}{\xi_n^2\sigma^2})\right)dx \tag{4.156}$$
$$> \frac{1}{2}\log 2\pi e+e^{-\frac{1}{2}}\frac{1}{\sqrt{2\pi}}\int_{0}^{1}\log(x^2+\frac{1}{\xi_n^2\sigma^2})dx+\log(1+\frac{1}{\xi_n^2\sigma^2})\int_{1}^{\infty}\frac{1}{\sqrt{2\pi}}e^{\frac{-x^2}{2}}dx \tag{4.157}$$
$$> \frac{1}{2}\log 2\pi e+\frac{e^{-\frac{1}{2}}}{\sqrt{2\pi}}\int_{0}^{1}\log(x^2+\frac{1}{\xi_n^2\sigma^2})dx+0.158\log(1+\frac{1}{\xi_n^2\sigma^2}) \tag{4.158}$$
$$= \frac{1}{2}\log 2\pi e+\frac{e^{-\frac{1}{2}}}{\sqrt{2\pi}}\left(x\left(\log(x^2+\frac{1}{\xi_n^2\sigma^2})-2\right)+\frac{2}{\sigma\xi_n}\tan^{-1}\sigma\xi_n x\right)\Bigg|_0^1+0.158\log(1+\frac{1}{\xi_n^2\sigma^2}) \tag{4.159}$$
$$= \frac{1}{2}\log 2\pi e+\frac{e^{-\frac{1}{2}}}{\sqrt{2\pi}}\left(1\cdot\left(\log(1+\frac{1}{\xi_n^2\sigma^2})-2\right)+\frac{2}{\sigma\xi_n}\tan^{-1}\sigma\xi_n-0\right)+0.158\log(1+\frac{1}{\xi_n^2\sigma^2}) \tag{4.160}$$
$$> \frac{1}{2}\log 2\pi e++\frac{e^{-\frac{1}{2}}}{\sqrt{2\pi}}\left(\log(1)-2+\frac{2}{\sigma\xi_n}\tan^{-1}\sigma\xi_n\right)+0.158\log(1) \tag{4.161}$$
$$> \frac{1}{2}\log 2\pi e+0.242\,(0-2+0)+0 \tag{4.162}$$
$$> \frac{1}{2}\log 2\pi e-0.48 \tag{4.163}$$

(4.158) follows since we are substituting a monotonic function by it's minimum value. (4.159) follows from Mathematica :). (4.162) follows since the inverse tangent of a positive number is positive. Hence,

$$I(X[n];Y[n]) < \frac{1}{2}\log\left(2\pi e(\frac{\mu^2}{\sigma^2}+1+\frac{1}{\xi_n^2\sigma^2})\right) - \frac{1}{2}\log 2\pi e + 0.48 \tag{4.164}$$

$$= \frac{1}{2}\log(\frac{\mu^2}{\sigma^2}+1+1) + 0.48 \tag{4.165}$$

$$= \frac{1}{2}\log(\frac{\mu^2}{\sigma^2}+2) + 0.48 \tag{4.166}$$

This bound goes to infinity as $\mu \to \infty$ and goes to 0 as with $\sigma \to \infty$ as the deterministic model would suggest. ■

Chapter 5

Control capacity

5.1 Introduction

This chapter uses an information-theoretic perspective to take a fresh look at stability in control systems. Communication capacity is used as a standard metric in information transmission systems. Previous results have shown that information flows and bottlenecks in systems clearly affect our ability to stabilize a system [110, 96]. These works treated the uncertainty in the plant as an information *source*. The data-rate theorems [110, 75, 80] are a product of using information-theoretic techniques to understand the impact of unreliable communication channels (connecting the sensor to the controller) on the ability to control a system. Anytime results [96] also focused on unreliable communication channels, but focused on reliability aspects of the sensing channel. This highlighted the differences between delay in control and communication: *when* a bit is learned matters in control systems, and the degree to which it matters depends on the sense of stability desired.

In addition to the source nature of information in control systems, we find there is also a *sink* nature to systems[1]. Control systems can reduce uncertainty about the world by moving the world to a known point: it is this dissipation of information/uncertainty that we refer to as the sink nature of a control system. To understand this further, this chapter focuses on the actuation channels in the system (Fig. 5.1), where the unreliability is about the control action and not the observation.

Information theory understands the capacity of a communication channel as the maximum number of bits of information that can be transmitted across the channel in a unit timestep. What is the maximum number of bits of the system state that can be dissipated by a controller in a single timestep? Is it important which bits are dissipated? It turns out the answer to the second question, in the context of control systems, is yes. To stabilize a system, the controller must dissipate the uncertainty in the most-significant bits (MSBs) of the state. We believe that the maximum number of these MSBs that can be dissipated is

[1]Work on Witsenhausen's counterexample shows that there is also a *channel* aspect to decentralized control systems [49, 83].

the “control capacity” of the system.

The work in this chapter is also motivated by ideas in portfolio theory. There, the key concept is the doubling rate of the system, i.e. the rate at which a gambler who chooses an optimal portfolio doubles his principal. Kelly studied this through bets placed on horse races in [60]. Control systems, like portfolios, have an underpinning of exponential growth. Just as the investor can choose to buy and sell at each time step to maximize growth, the controller has the choice of control strategy to minimize growth (or maximize decay). The control capacity of a system is closely related to the “tolerable growth rate” for the system. A larger control capacity means that the system can tolerate a higher intrinsic rate of uncertainty growth[2]. This also parallels the classic sense in communication channels, where the rate is how fast the set of potential messages grows with blocklength.

There are two big advantages to the information-theoretically motivated definitions of control capacity here. First, they more easily allow us to measure the impact of side information to improve performance in systems. This is discussed further in Chapter 6. Second, they allow us to move beyond second-moment notions of state stability. We define a range of notions of capacity corresponding to different moments of stability. The weakest notion of stability is given by the logarithmic decay rate of the system, and we call the related capacity notion “Shannon” control capacity. This is inspired by the results on the anytime capacity of a channel, where it turns out that as the desired sense of stability for the system goes to zero, the anytime capacity converges to the Shannon capacity of the corresponding communication channel [96]. The strictest notion is of control capacity is zero-error control capacity, which is similar to the worst-case perspective of robust control. The squared-error notion of stability lies in between these two.

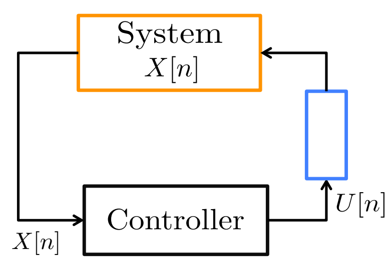


Figure 5.1: This chapter focuses on the control limitations due to uncertainties on the actuation channel in the system.

Unreliable actuation channels may arise in the context of high-reliability low-latency wireless control systems, where dumb actuators receive control signals over packet-dropping networks or fast-fading channels. Bio-medical technologies such as brain-machine interfaces will have to rely on wirelessly connected low-power unreliable actuators and sensors, and

[2]Results in Chapter 6 motivated the definitions of control capacity here. Tolerable growth is discussed in more detail there.

moreover the biological actuation system its own eccentricities. These issues also emerge in the development of "wearable" technologies, and more broadly speaking in the development of the Internet of Things (IoT). This chapter takes a look at simplified models with i.i.d. randomness to get a theoretical understanding of the basic issues.

5.2 Actuation channels

We start by exploring the second-moment stability in a simple example that motivates the work.

5.2.1 Second-moment stability

First, we recall the definition of mean-square stability that is the most widely used notion of stability in the control community today.

Definition 3.1.2. Let the evolution of system state $X[n]$ be governed by a linear system

$$\begin{aligned} X[n+1] &= A[n]\cdot X[n] + B[n]\cdot U[n] + W[n] \\ Y[n] &= C[n]\cdot X[n] + V[n] \end{aligned} \tag{3.1}$$

$A[n], B[n], C[n], V[n], W[n]$ are random variables with known distributions. Let $\sigma(Y_0^n)$ be the sigma-algebra generated by the random variables $Y[0]$ to $Y[n]$. The system is said to be **mean-square stabilizable** if there exists a causal control strategy $U_0^n(\cdot) \in \sigma(Y_0^n)$ such that $X[n]$ is mean-square stable, i.e. there exists $M \in \mathbb{R}$, $M < \infty$, s.t. $\mathbb{E}[|X[n]|^2] < M$ for all n.

Consider a simple scalar control system with perfect state observation, and no additive driving disturbance, as below. This is a special case of the uncertainty threshold principle setup [6] discussed in Chapter 3.

$$\begin{aligned} X[n+1] &= aX[n] + B[n]U[n], \\ Y[n] &= X[n]. \end{aligned} \tag{5.1}$$

Suppose $B[n]$ are a sequence of i.i.d. random variables with mean μ_B and variance σ_B^2, and $X[0] \sim \mathcal{N}(0,1)$. The system state is scaled up by a scalar constant factor a at each time step. The aim is to choose $U[n]$, a function of $Y[n]$, so as to stabilize the system.

Example 5.2.1. System (5.1) is mean-square stabilizable using linear strategies if $a^2 < \left(1 + \frac{\mu_B^2}{\sigma_B^2}\right)$.

Let $U[n] = k \cdot Y[n] = k \cdot X[n]$, where k is a scalar constant. Then,

$$\mathbb{E}[X[n+1]^2] = \mathbb{E}[(a + kB[n])^2 X[n]^2] \tag{5.2}$$

$$= (a^2 + 2ak\mu_B + k^2(\mu_B^2 + \sigma_B^2))\mathbb{E}[X[n]^2] \tag{5.3}$$

The scale factor, $(a^2 + 2ak\mu_B + k^2(\mu_B^2 + \sigma_B^2))$, is minimized at $k = -\frac{a\mu_B}{\mu_B^2+\sigma_B^2}$. Thus, with a linear strategy, the minimum growth rate is given by $a^2\frac{\sigma_B^2}{\mu_B^2+\sigma_B^2}$, which is smaller than 1 if $a^2 < \left(\frac{\mu_B^2+\sigma_B^2}{\sigma_B^2}\right)$. ■

The uncertainty threshold principle gives the achievable strategy and converse for this system for mean-squared error and shows that linear strategies are optimal in this case. The threshold here shows that the randomness in the control parameter $B[n]$ is the "bottleneck" on the controller's ability to stabilize the system. To tolerate a larger system growth a we need to somehow reduce the randomness in $B[n]$, or change what we mean by stability.

5.2.2 Logarithmic stability

We would also like to consider log-scale stability of systems, in order to capture a notion of growth rate as in portfolio theory. Portfolio theory and information theory generally suggest looking at the relevant expectations of logs, to find the natural growth rate of the system. We consider the logarithm of the system magnitude and aim to maintain $\lim_{n\to\infty} \mathbb{E}\ \log |X[n]| < \infty$. Log-scale stability is implied by mean-squared stability, since the log of a positive real number is strictly smaller than its square. Now we define a logarithmically stable system.

Definition 5.2.2. Let the evolution of system state $X[n]$ be governed by a linear system

$$\begin{aligned} X[n+1] &= A[n] \cdot X[n] + B[n] \cdot U[n] + W[n] \\ Y[n] &= C[n] \cdot X[n] + V[n] \end{aligned} \tag{5.4}$$

$A[n], B[n], C[n], V[n], W[n]$ are random variables with known distributions at each time n. Let $\sigma(Y_0^n)$ be the sigma-algebra generated by the random variables $Y[0]$ to $Y[n]$. The system is said to be **logarithmically stabilizable** if there exists a causal control strategy $U_0^n(\cdot) \in \sigma(Y_0^n)$ such that $\exists M \in \mathbb{R}$, $M < \infty$, s.t. $\mathbb{E}[\log |X[n]|] < M$ for all n.

Remark 5.2.3. Note that all logarithms in this chapter are taken to the base 2.

Example 5.2.4. Consider the system (5.1) with $B[n] \sim \text{Uniform}[b_1, b_2]$, where $a, b_1, b_2 > 0$, and use a memoryless linear control strategy $U[n] = s \cdot X[n]$.

Choose a $U[n] = s \cdot X[n]$. To minimize the expected logarithmic growth rate, we want to minimize $\mathbb{E}\log\left|\frac{X[n+1]}{X[n]}\right| = \mathbb{E}\log|a(1 + B[n]s)|$. The choice $B[n] \sim \text{Uniform}[b_1, b_2]$ allows for explicit evaluation of the integral that gives the expectation, i.e.

$$\int_{b_1}^{b_2} \log |a(1 + b \cdot s)| \frac{1}{b_2 - b_1} db. \tag{5.5}$$

Consider,

$$\frac{\partial}{\partial s}\int_{b_1}^{b_2} \log|a(1+b\cdot s)|\frac{1}{b_2-b_1}db \tag{5.6}$$

$$= \int_{b_1}^{b_2} \frac{\partial}{\partial s}\log|a(1+b\cdot s)|\frac{1}{b_2-b_1}db \tag{5.7}$$

$$= \int_{b_1}^{b_2} \frac{b}{|1+b\cdot s|}sgn(1+b\cdot s)\frac{1}{b_2-b_1}db \tag{5.8}$$

$$= \int_{b_1}^{b_2} \frac{b}{(1+b\cdot s)}\frac{1}{b_2-b_1}db \tag{5.9}$$

$$= \frac{1}{b_2-b_1}\frac{b\cdot s-\log(1+b\cdot s)}{s^2}\Bigg|_{b_1}^{b_2} \tag{5.10}$$

$$= \frac{1}{s}+\frac{1}{s^2(b_2-b_1)}\log\left|\frac{1+b_2\cdot s}{1+b_1\cdot s}\right| \tag{5.11}$$

Setting this equal to zero would give the minimizing s, however the equations are non-algebraic, so the minimizing control cannot be found analytically for general b_1 and b_2. However, it can be easily computed. To gain some intuition we can choose the slightly suboptimal strategy $s = -\frac{1}{\frac{b_1+b_2}{2}}$. This gives the expected logarithmic system growth as $\log|a| + \mathbb{E}\log|1-\frac{2B[n]}{b_1+b_2}|$, and this strategy is successful if the quantity is negative. Reducing the variance of $B[n]$ through side information would improve the performance of this strategy. However, note that this $U[n]$ is *not* the optimal control strategy.

Again, the stability threshold here is determined by the uncertainty in $B[n]$. But is this a real bottleneck or just something artificial due to our restriction to linear memoryless strategies? We need a converse that says we cannot do any better. For both of these examples side information about continuous uncertainty on the control gain is useful, in both a mean-squared and logarithmic sense. This leads into the developments in Chapter 6 where we can explicitly incorporate side information.

Remark 5.2.5. We can also consider the example where the $B[i]$'s are i.i.d. Gaussian random variables as in . That case is similar to the Uniform case in that we are limited to numerically computing the optimal control there as well. We include the Uniform example because we can evaluate the suboptimal strategy of "aiming for the mean" easily. We will see that while this is a good strategy in the robust control or "zero-error" sense of stability, it is not good in logarithmic sense. Plots at the end of this chapter will show how these different senses of stability lead to different optimal strategies in both the Gaussian and Uniform case.

5.3 Carry-free systems

We would like to finally move to define the control capacity of a system.

As always, we start with carry-free models to understand the problem. We focus on a system with only actuation uncertainty.

Consider the carry-free system, $\mathcal{S}$, with a random gain for the control input, with unit system gain (Fig. 5.2).

$$\begin{aligned} x[n+1](z) &= x[n](z) + b[n](z) \cdot u[n](z) + w[n](z) \quad (5.12)\\ y[n](z) &= x[n](z). \quad (5.13) \end{aligned}$$

and

$$b[n](z) = 1 \cdot z^{g_{det}} + 0 \cdot z^{g_{det}-1} + 0 \cdot z^{g_{det}-2} + \cdots + b_{g_{ran}}[n] \cdot z^{g_{ran}} + b_{g_{ran}-1}[n] \cdot z^{g_{ran}-1} + \cdots \quad (5.14)$$

Basically, b is a 1, followed by 0's, followed by random numbers. There are $g_{det} - g_{ran}$ deterministic bits in it. As before we have:

$$g_{det} = \begin{cases} \max\{i \mid b_i = 1\} \text{ if } g_{ran} < \text{ degree of } b[\cdot](z) \\ g_{ran}, \text{ otherwise.} \end{cases} \quad (5.15)$$

$w[n](z)$ is similar to $b[n](z)$ with all bits being random, and degree upperbounded by 0.

Let the degree of $x[n]$ be denoted as d_n. Our aim is to understand the stability of this system, which is captured by the degree d_n. First, we define the notion of "in-the-box" or zero-error stability, where the degree of the state remains bounded with probability 1.

Definition 5.3.1. The system eq. (5.12) is stablizable in the zero-error sense if there exists a control strategy $u[\cdot](z)$ such that $\exists N, M < \infty,$ s.t. $\forall n > N, d_n < M$ with probability 1.

What about the average case? For this we define the "Shannon" sense of stability. The reason behind the terminology will become clear later.

Definition 5.3.2. The system eq. (5.12) is stablizable in the "Shannon" sense if there exists a control strategy $u[\cdot](z)$ such that $\exists M < \infty$ s.t. $\limsup_{n\to\infty} \mathbb{E}[d_n] < M$.

With these definitions, the picture clearly illustrates the ideas that we want to capture with control capacity. How many bits-levels can we reduce the degree by with probability 1? This should be the zero-error capacity, which we define formally as follows.

Definition 5.3.3. The zero-error control capacity of the system $\mathcal{S}$ from eq. (5.12) is defined as the largest constant $C_{ze}(\mathcal{S})$ such that there exists a control strategy $u[0](z), \ldots, u[n](z)$ such that

$$P\left(\frac{1}{n}(d_0 - d_n) \geq C_{ze}(\mathcal{S})\right) = 1. \quad (5.16)$$

The time index n does not matter.

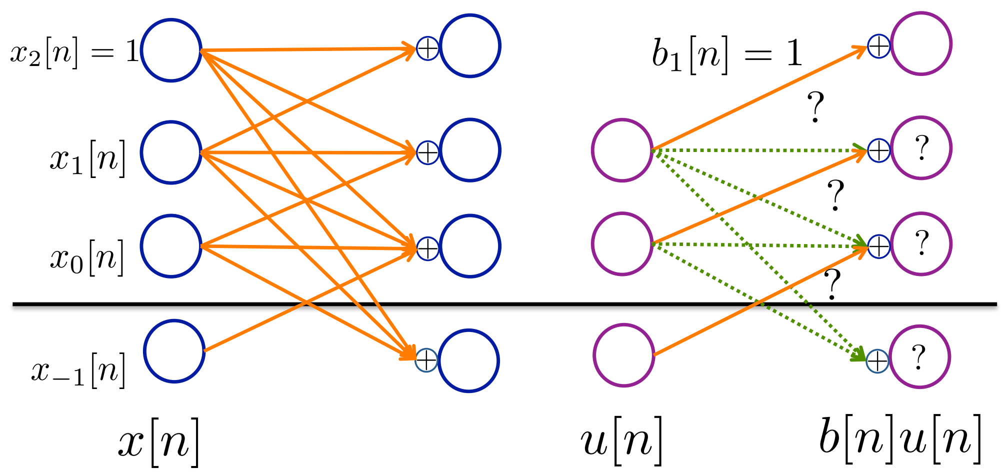


Figure 5.2: A carry-free control model with varying control gain but system gain 1 (Eq. (5.12)).

C_{ze} is essentially the largest decay exponent that is possible for the system state.

On the other hand, we might ask, what is the expected decrease in degree? This is the "Shannon" control capacity.

Definition 5.3.4. The "Shannon" control capacity of the system $\mathcal{S}$ from eq. (5.12), $C_{sh}(\mathcal{S})$, is defined as $\lim_{n\to\infty} \frac{1}{n}\mathbb{E}[d_0 - d_n]$ under the best choice of control strategy $u[0](z), \ldots, u[n](z)$.

Now that we understand this for a system with no system gain $a(z)$, it is clear how to extend to non-trivial gains. The two notions of capacity above clearly relate to two different notions of stability. The "Shannon" notion of stability is weaker than the zero-error notion. In both cases the degree-gain of the system needs to be smaller than the number of bits that can be cancelled.

Theorem 5.3.5. *Consider the system $\widetilde{S}^a$:*

$$\widetilde{x}[n+1](z) = a(z)(\widetilde{x}[n](z) + b[n](z) \cdot \widetilde{u}[n](z) + w[n](z)), \tag{5.17}$$

where $a(z)$ is a constant known polynomial with gain g_a, as defined in Chapter 3.

Also consider the affiliated system $\mathcal{S}$ defined above in eq. (5.12). Then,

- $\exists N, M < \infty$, *s.t.* $\forall n > N, d_n < M$ *with probability* 1 *if and only if* $C_{ze}(\mathcal{S}) \geq g_a$.
- $\lim_{n\to\infty} \mathbb{E}[d_n] < \infty$ *if* $C_{sh}(\mathcal{S}) > g_a$. $\mathbb{E}[d_n]$ *grows unboundedly if* $C_{sh}(\mathcal{S}) < g_a$.

Proof: The proof follows naturally from the definitions. ∎

This shows us that the control capacity only depends on system through the actuation channel, i.e. the uncertainty in $b[n]$. We will see parallel results in the real case.

5.3.1 Zero-error control capacity

Theorem 5.3.6. *The zero-error control capacity for the system defined by eq. (5.12) is given by*

$$C_{ze}\left(\mathcal{S}\right) = g_{det} - g_{ran}. \tag{5.18}$$

This theorem is essentially a restatement of Thm. 3.4.5, the carry-free uncertainty threshold principle for unknown control gains from Chapter 3.

Before we prove this theorem, here is a key lemma that bounds the decay that can happen in one step, regardless of the system state $x[n]$.

Lemma 5.3.7. *For the system defined by eq. (5.12), for* ***any*** *state $x[n]$ the largest constant $C_{ze,n}$ such that $P\left(d_n - d_{n+1} \geq C_{ze,n}\right) = 1$ is $C_{ze,n} = g_{det} - g_{ran}$,*

Proof:
Achievability: The achievability follows naturally by solving the appropriate set of linear equations to calculate controls to cancel the bits of the state.

Converse: To show the converse, we must show that for any $x[n]$ and for $u[n]$ that depends on $x[n]$ and its history we cannot beat $g_{det} - g_{rand}$,

$$P\left(d_n - d_{n+1} < g_{det} - g_{ran} + 1\right) > 0. \tag{5.19}$$

Consider any $x[n] = x_{d_n} z^{d_n} + x_{d_n-1} z^{d_n-1} + \ldots$, with degree d_n. Then, $x_{d_n} = 1$.
Choose any $u[n] = u_{m_n} z^{m_n} + u_{m_n-1} z^{m_n-1} + \ldots$.

$$\begin{aligned} x[n+1] =& (x_{d_n}[n] z^{d_n} + x_{d_n-1}[n] z^{d_n-1} + \ldots) + \\ & (z^{g_{det}} + b_{g_{ran}}[n] z^{g_{ran}} + b_{g_{ran}-1}[n] \cdot z^{g_{ran}-1} + \ldots) \cdot (u_{m_n}[n] z^{m_n} + u_{m_n-1}[n] z^{m_n-1} + \ldots) \end{aligned} \tag{5.20}$$

Note $g_{det} - g_{ran} \geq 0$ by definition of g_{det} and g_{ran}.

- Case 1) $g_{det} + m_n > d_n$

  In this case, $d_{n+1} = g_{det} + m_n$. So

  $$d_n - d_{n+1} = d_n - (g_{det} + m_n) \leq 0 < g_{det} - g_{ran} + 1. \tag{5.21}$$

- Case 2) $g_{det} + m_n < d_n$

  In this case, $d_{n+1} = d_n$. So

  $$d_n - d_{n+1} = 0 < g_{det} - g_{ran} + 1. \tag{5.22}$$

- Case 3)$g_{det} + m_n = d_n$

  We know $u_{m_n}[n] = 1$.

  Now, consider the coefficient of $z^{d_n - g_{ran} + g_{det}}$ in $x[n+1]$:

  $$x_{d_n - g_{ran} + g_{det}}[n] + b_{g_{ran}}[n] \cdot u_{m_n}[n] + b_{g_{det}}[n] \cdot u_{g_{ran}+d_n-2g_{det}}[n] \tag{5.23}$$

  To see where the second and third terms come from, remember that $g_{det} + m_n = d_n$, and all coefficients of b between $z^{g_{det}}$ and $z^{g_{ran}}$ are zero.

  Consider the term below. Recall $b_{g_{set}}[n] = 1$.

  $$\begin{aligned} &x_{d_n - g_{ran} + g_{det}}[n] + b_{g_{ran}}[n] \cdot 1 + 1 \cdot u_{g_{ran}+d_n-2g_{det}}[n] \\ =&x_{d_n - g_{ran} + g_{det}}[n] + b_{g_{ran}}[n] + u_{g_{ran}+d_n-2g_{det}}[n] \end{aligned} \tag{5.24}$$

  Since here $b_{g_{ran}}[n]$ is a Bernoulli$-(\frac{1}{2})$, this term will be zero exactly with probability $\frac{1}{2}$.

  Hence, with probability $\frac{1}{2}$, $d_{n+1} \geq d_n - g_{ran} + g_{det}$. Hence, with probability $\frac{1}{2}$, $d_n - d_{n+1} \leq g_{ran} - g_{det}$. Thus $P(d_n - d_{n+1} < 1 + g_{ran} - g_{det}) \geq \frac{1}{2}$, which gives the converse.

■

With Lemma 5.3.7 we can prove Thm. 5.3.6. This proof follows easily since the lemma, in effect, decouples the controls at different time steps. We are thus freed from considering time-varying or state-history dependent control strategies, which generally makes this style of converse difficult.

Proof: (Thm. 5.3.6)

The lemma above bounds the decrease in degree of the state at any given time n, regardless of the control $u[n](z)$ or the state of the system $x[n](z)$.

$$P\left(\frac{1}{n}(d_0 - d_n) > C_{ze}(\mathcal{S})\right) = 1. \tag{5.25}$$

$$\Longleftrightarrow P\left(\frac{1}{n}\sum_{i=0}^{n-1}(d_i - d_{i+1}) > C_{ze}(\mathcal{S})\right) = 1. \tag{5.26}$$

We know from Lemma 5.3.7 that,

$$P\left(\frac{1}{n}\sum_{i=0}^{n-1}(d_i - d_{i+1}) > \sum_{i=0}^{n-1}(g_{det} - g_{ran})\right) = 1. \tag{5.27}$$

Hence, we must have

$$C_{ze}(\mathcal{S}) = \frac{1}{n}\sum_{i=0}^{n-1}(g_{det} - g_{ran}) = g_{det} - g_{ran}. \tag{5.28}$$

■

5.3.2 "Shannon" control capacity

We can similarly calculate the Shannon control capacity. The proof follows the same pattern as the proof for zero-error control capacity.

Theorem 5.3.8. *The Shannon control capacity for the system $\mathcal{S}$ is given by*

$$C_{sh}(\mathcal{S}) = g_{det} - g_{ran} + 1. \tag{5.29}$$

Lemma 5.3.9. *For the system defined by eq. (5.12), for **any** state $x[n]$ history,* $\max_{u[n]} \mathbb{E}[d_n - d_{n+1}] = g_{det} - g_{ran} + 1$.

Proof:
Achievability: Just as in the zero-error case, we can solve a system of linear equations to get the right achievable strategy. Considering expectations gives the extra bit because the initial run of zeros we can expect in the random part is basically a geometric random variable with expectation 1. Consequently, we can apply a control pretending that the random part of the B is all 0s and on average, get one bit of benefit from doing so.
Converse: For the converse, we need to show that $\max_{u[n]} \mathbb{E}[d_n - d_{n+1}] \leq g_{det} - g_{ran} + 1$. Let $x[n] = x_{d_n} z^{d_n} + u_{d_n-1} z^{d_n-1} + \cdots$, and let $x_{d_n} = 1$. Consider any control with a fixed non-random degree m_n, $u[n] = u_{m_n} z^{m_n} + u_{m_n-1} z^{m_n-1} + \cdots$.

Then,

$$\begin{aligned} x[n+1] =& (x_{d_n}[n] z^{d_n} + x_{d_n-1}[n] z^{d_n-1} + \cdots) + \\ & (z^{g_{det}} + b_{g_{ran}}[n] z^{g_{ran}} + \cdots) \cdot (u_{m_n}[n] z^{m_n} + u_{m_n-1}[n] z^{m_n-1} + \cdots) \end{aligned} \tag{5.30}$$

Note $g_{det} - g_{ran} \geq 0$ by definition of g_{det} and g_{ran}.

- Case 1) $g_{det} - g_{ran} > 0$ and $g_{det} + m_n > d_n$
  In this case, $d_{n+1} = g_{det} + m_n$. So

  $$\mathbb{E}[d_n - d_{n+1}] = \mathbb{E}[d_n - g_{det} + m_n] \leq 0 \leq g_{det} - g_{ran} + 1. \tag{5.31}$$

  since

- Case 2) $g_{det} - g_{ran} > 0$ and $g_{det} + m_n < d_n$ In this case, $d_{n+1} = d_n$. So

  $$\mathbb{E}[d_n - d_{n+1}] = 0 \leq g_{det} - g_{ran} + 1. \tag{5.32}$$

- Case 3) $g_{det} - g_{ran} > 0$ and $g_{det} + m_n = d_n$
  So, $u_{m_n}[n] = 1$.

Now, consider the coefficient of $z^{d_n - g_{det} + g_{ran}}$ in $x[n+1]$:

$$x_{d_n - g_{det} + g_{ran}}[n] + b_{g_{ran}}[n] \cdot u_{m_n}[n] + b_{g_{det}}[n] \cdot u_{d_n - 2g_{det} + g_{ran}}[n] \tag{5.33}$$

To see where the second and third terms come from remember that $g_{det} + m_n = d_n$, and all coefficients of b between $z^{g_{det}}$ and $z^{g_{ran}}$ are zero.

Consider the term below. Recall $b_{g_{det}}[n] = 1$.

$$\begin{aligned} &x_{d_n + g_{ran} - g_{det}}[n] + b_{g_{ran}}[n] \cdot 1 + 1 \cdot u_{d_n + g_{ran} - 2g_{det}}[n] \\ =&x_{d_n + g_{ran} - g_{det}}[n] + b_{g_{ran}}[n] + u_{g_{ran} + d_n - 2g_{det}}[n] \end{aligned} \tag{5.34}$$

Since here $b_{g_{ran}}[n]$ is a Bernoulli$-(\frac{1}{2})$, this term will be zero exactly with probability $\frac{1}{2}$.

Now, consider the coefficient of $z^{d_n - g_{ran} + g_{det} - k}$ in x[n+1] for $k > 0$:

$$x_{d_n + g_{ran} - g_{det} - k}[n] + b_{g_{det}}[n] \cdot u_{g_{ran} + d_n - 2g_{det} - k}[n] + \sum_{i=0}^{k} b_{g_{ran} - i}[n] \cdot u_{m_n - k + i}[n] \tag{5.35}$$

$$= x_{d_n + g_{ran} - g_{det} - k}[n] + u_{g_{ran} + d_n - 2g_{det} - k}[n] + \sum_{i=0}^{k} b_{g_{ran} - i}[n] \cdot u_{m_n - k + i}[n] \tag{5.36}$$

$$= x_{d_n + g_{ran} - g_{det} - k}[n] + u_{g_{ran} + d_n - 2g_{det} - k}[n] + b_{g_{ran} - k}[n] \cdot u_{m_n}[n] \sum_{i=0}^{k-1} b_{g_{ran} - i}[n] \cdot u_{m_n - k + i}[n] \tag{5.37}$$

The third term here is a Bernoulli$-(\frac{1}{2})$ independent of other terms, since $b_{g_{ran} - k}$'s are independent for every k. So this coefficient is 0 with probability $\frac{1}{2}$.

Hence, $P(d_{n+1} = d_n + g_{ran} - g_{det} - k) = \frac{1}{2^{k+1}}$. So then

$$\begin{aligned} \mathbb{E}[d_n - d_{n+1}] &\leq \sum_{k=0}^{\infty} (g_{det} - g_{ran} + k) \frac{1}{2^{k+1}} && (5.38) \\ &= g_{det} - g_{ran} + 1 && (5.39) \end{aligned}$$

- Finally, Case 0) $g_{det} - g_{ran} = 0$
  If $m_n + g_{ran} < d_n$, then $d_{n+1} = d_n$ and we are done.
  $m_n + g_{ran} > d_n$ is clearly suboptimal and by the argument in Case (1) we are done.
  If $m_n + g_{ran} = d_n$, then by arguments similar to case 3, $P(d_{n+1} = m_n + g_{ran} - k) = \frac{1}{2}^{k+1}$ for every m_n. This is true since the highest possible level for $b[n]u[n]$ is $g_{ran} + m_n$.

Hence,

$$\mathbb{E}[d_n - d_{n+1}] = d_n - \sum_{k=0}^{\infty}(m_n + g_{ran} - k) \tag{5.40}$$

$$\leq d_n - \sum_{k=0}^{\infty}(d_n - k) \tag{5.41}$$

$$= d_n - (d_n - 1) \tag{5.42}$$

$$= 1 \tag{5.43}$$

which proves the result.

■

Now we can prove 5.3.8.

Proof: (Thm. 5.3.8) The argument here is essentially identical to the one made in the zero-error carry-free control capacity case. The fact that Lemma 5.3.9 holds for all states $x[n]$ again plays a key role. ■

With this, we move to consider real-valued systems, starting with the zero-error notions of stability and control capacity.

5.4 Zero-error control capacity

Definition 5.4.1. Let the evolution of system state $X[n]$ be governed by a linear system

$$X[n+1] = A[n] \cdot X[n] + B[n] \cdot U[n] + W[n]$$
$$Y[n] = C[n] \cdot X[n] + V[n] \tag{5.44}$$

$A[n], B[n], C[n], V[n], W[n]$ are random variables with known distributions at each time n. Let $\sigma(Y_0^n)$ be the sigma-algebra generated by the random variables $Y[0]$ to $Y[n]$. The system is said to be stabilizable in the **zero-error** sense if $\exists N, M < \infty$ and a causal control strategy U_0^∞ such that $U[i](\cdot) \in \sigma(Y_0^i)$ $\forall i$, such that $\forall n > N$,

$$P(|X[n]| < M) = 1. \tag{5.45}$$

Consider the following system $\mathcal{S}(p_{B[0]}, p_{B[1]}, \ldots)$, with no additive noise V, or additive disturbance W, and with system dynamics $A[n] = 1$.

$$X[n+1] = X[n] + B[n]U[n]$$
$$Y[n] = X[n] \tag{5.46}$$

The control signal $U[n]$ can causally depend on $Y[i], 0 \leq i \leq n$. Let $\sigma(Y_0^n)$ be the sigma-algebra generated by the observations. Then $U[n]$ can be a function of these random variables, i.e. $U[n] \in \sigma(Y_0^n)$. The random variables $B[i], 0 \leq i \leq n$ are independent, and

$B[i] \sim p_{B[i]}$. The distribution of each $B[i]$ is known to the controller beforehand. We will consider different restrictions on the initial state $X[0]$ later, but in general, we consider $X[0]$ to be a random variable with density $p_{X[0]}(\cdot)$ and no atom at 0.

Definition 5.4.2. The zero-error control capacity of the system $\mathcal{S}(p_{B[0]}, p_{B[1]}, \cdots)$ is defined as

$$C_{ze}(\mathcal{S}) = \max\left\{C \middle| \exists N, \text{ s.t. } \forall n > N, \max_{U_1^n \text{ s.t. } U[i](\cdot)\in\sigma(Y_0^i)} P\left(-\frac{1}{n}\log\left|\frac{X[n]}{X[0]}\right| > C\right) = 1\right\}. \tag{5.47}$$

This is a well defined quantity since $p_{X[0]}(\cdot)$ has no atoms at 0. It could be infinite however.

Consider also the system $\widetilde{\mathcal{S}}^a(p_{B[0]}, p_{B[1]}, \cdots)$, with no additive disturbances and noise but with non-trivial intrinsic growth a,

$$\begin{aligned} \widetilde{X}[n+1] &= a \cdot (\widetilde{X}[n] + B[n]\widetilde{U}[n]) \\ \widetilde{Y}[n] &= \widetilde{X}[n] \end{aligned} \tag{5.48}$$

with initial condition $\widetilde{X}[0]$ identical to system $\mathcal{S}$, and such that $p_{\widetilde{X}[0]}(\cdot)$ has bounded support and no atom at 0. Again, $\widetilde{U}[n] \in \sigma(\widetilde{Y}_0^n)$. The random variables $B[i]$ in the system have the same distributions as those of system $\mathcal{S}$.

Theorem 5.4.3. *If the zero-error control capacity of the system $\mathcal{S}(p_{B[0]}, p_{B[1]}, \cdots)$ (eq. (5.46)),*

$$C_{ze}(\mathcal{S}) \geq \log|a|, \tag{5.49}$$

then the associated system $\widetilde{\mathcal{S}}^a(p_{B[0]}, p_{B[1]}, \cdots)$, with the identical initial condition $\widetilde{X}[0] = X[0]$, such that $p_{X[0]}(\cdot)$ has bounded support, and with the same distributions for the random control gains $B[i]$, is stabilizable in a zero-error sense.

Further, if the system $\widetilde{\mathcal{S}}^a(p_{B[0]}, p_{B[1]}, \cdots)$ is stabilizable in a zero-error sense, then the zero-error control capacity of the system $\mathcal{S}(p_{B[0]}, p_{B[1]}, \cdots)$ satisfies

$$\forall \epsilon > 0, C_{ze}(\mathcal{S}) \geq \log|a| - \epsilon, \tag{5.50}$$

The fact that this theorem is true, namely the fact that the zero-error control capacities for system $\mathcal{S}(p_{B[0]}, p_{B[1]}, \cdots)$ lets us answer a question about the system $\widetilde{\mathcal{S}}^a(p_{B[0]}, p_{B[1]}, \cdots)$, justifies thinking of the control capacity as a property of the actuation channel.

Before we prove the theorem, here is a useful lemma, that shows that the two systems, $\mathcal{S}(p_{B[0]}, p_{B[1]}, \cdots)$ and $\widetilde{\mathcal{S}}^a(p_{B[0]}, p_{B[1]}, \cdots)$, can be made to track each other.

Lemma 5.4.4. *Let U_0^∞ be a set of controls applied to $\mathcal{S}(p_{B[0]}, p_{B[1]}, \cdots)$. Let $\widetilde{U}[i] = a^i U[i]$ be the controls applied to $\widetilde{\mathcal{S}}^a(p_{B[0]}, p_{B[1]}, \cdots)$, with the same realizations of the random variables $B[i]$ and initial condition $X[0]$. Then, $\widetilde{U}[i](\cdot) \in \sigma(\widetilde{Y}_0^i)$ and $\forall\ i, \widetilde{X}[i] = a^i X[i]$.*

Proof: The proof proceeds by induction. $X[0] = \widetilde{X}[0]$ serves as the base case. Further, $\widetilde{U}[0] \in \sigma(\widetilde{Y}[0])$, since $\widetilde{X}[0] = X[0]$.

$$
\begin{aligned}
\widetilde{X}[i+1] &= a \cdot (\widetilde{X}[i] + B[i]\widetilde{U}[i]) && (5.51)\\
&= a \cdot (a^i \widetilde{X}[i] + a^i B[i]\widetilde{U}[i]) && (5.52)\\
&\text{(since } \widetilde{Y}[j] = a^j Y[j] \ \forall j \le i,\ \widetilde{U}[i](\cdot) \in \sigma(\widetilde{Y}_0^i))\\
&= a^{i+1} \cdot (X[i] + B[i]U[i]) && (5.53)\\
&= a^{i+1} X[i+1] && (5.54)
\end{aligned}
$$

■

Proof: **(Thm. 5.4.3)** ($\Longrightarrow$) We start with the first part of the theorem, to show stabilizability with high enough control capacity. We know $\exists N,$ s.t. $\forall n > N$, $\exists U_1^n$ s.t. $U[i](\cdot) \in \sigma(Y_0^i)$, and

$$
\begin{aligned}
1 =& P\left(-\frac{1}{n}\log\left|\frac{X[n]}{X[0]}\right| \ge \log|a|\right) && (5.55)\\
=& P\left(\left|\frac{X[n]}{X[0]}\right| \le 2^{-n\log|a|}\right) && (5.56)\\
=& P\left(|X[n]| \le |a|^{-n} \cdot |X[0]|\right) && (5.57)\\
=& P\left(|a|^n \cdot |X[n]| \le |X[0]|\right) && (5.58)\\
=& P\left(|\widetilde{X}[n]| \le |X[0]|\right) \ \text{(apply the controls generated by Lemma 5.4.4.)} && (5.59)
\end{aligned}
$$

Since $|X[0]|$ is bounded, choose $M = |X[0]|$ as the bound.
($\Longleftarrow$) Next, we assume the stabilizability of $\widetilde{\mathcal{S}}$. Then, we know $\exists N, M,$ s.t. $\forall n > N$, $\exists \widetilde{U}_1^n$ s.t. $\widetilde{U}[i] \in \sigma(\widetilde{Y}_0^i)$, and

$$
\begin{aligned}
1 =& P\left(|\widetilde{X}[n]| < M\right) && (5.60)\\
=& P\left(|a|^n \cdot |X[n]| < M\right) \ \text{(apply the controls generated by Lemma 5.4.4.)} && (5.61)\\
=& P\left(|X[n]| < |a|^{-n} \cdot M\right) && (5.62)\\
=& P\left(\left|\frac{X[n]}{X[0]}\right| < 2^{-n\log|a|}\frac{M}{|X[0]|}\right) \ \text{(}|X[0]|\text{ has no atom at 0)} && (5.63)\\
=& P\left(-\frac{1}{n}\log\left|\frac{X[n]}{X[0]}\right| > \log|a| - \frac{1}{n}\frac{M}{|X[0]|}\right) && (5.64)
\end{aligned}
$$

Now, for any $\epsilon > 0, \exists N_1$ such that $\forall n \ge N_1, \frac{1}{n}\frac{M}{|X[0]|} < \epsilon$. Hence, for $n \ge \max\{N, N_1\}$, we have

$$
-\frac{1}{n}\log\left|\frac{X[n]}{X[0]}\right| \ge \log|a| - \epsilon. \tag{5.65}
$$

■

5.4.1 Calculating zero-error control capacity

The operational definitions for zero-error control capacity involve an optimization over an infinite sequence of potential control laws. Here we show that in fact this quantity can be easily computed in the case where the $B[n]$'s are i.i.d. over time. First, we show that if $p_B(\cdot)$ does not have bounded support, the zero-error control capacity is 0.

Theorem 5.4.5. *The zero-error control capacity of the system eq. (5.46), $\mathcal{S}(p_B)$ (with i.i.d. $B[n]$'s and hence parameterized by the single distribution $p_B(\cdot)$), where $p_B(\cdot)$ does not have bounded support is zero.*

Proof: Consider any strategy $U_0^{n-1}(\cdot)$. Then, for every value of $X[n-1]$ and chosen $U[n-1] \neq 0$ there exists a realization of $B[n]$ that just pushes it over any boundary, i.e., $\forall\ M < \infty$, $P\left(|X[n]| > M\right) > 0$. If all $U[i] = 0$ for $0 \leq i \leq n-1$ then the capacity is zero trivially, since $X[0] = X[n]$. ■

Before we are able to calculate C_{ze} for general distributions with bounded support, we introduce a few lemmas.

Lemma 5.4.6.

$$\min_{k \in \mathbb{R}} \max_{B \in [b_1, b_2]} \left|1 + B \cdot k\right| = \begin{cases} \left|\frac{b_2 - b_1}{b_2 + b_1}\right| & \text{if } \ 0 \notin [b_1, b_2], \ \text{i.e. } 0 < b_1 < b_2 \text{ or } b_1 < b_2 < 0. \\ 1 \quad \text{if } \ 0 \in [b_1, b_2] \ \text{i.e. } b_1 \leq 0 \text{ and } b_2 > 0, \ \text{or } b_1 < 0 \text{ and } b_2 \geq 0. \end{cases} \tag{5.66}$$

This implies that if B is a random variable with essential support $[b_1, b_2]$, essential infimum b_1, essential supremum b_2, then $\forall \epsilon > 0$,

$$P\left(\log\left|1 + B \cdot k\right| < \log\left|\frac{b_2 - b_1}{b_1 + b_2}\right| - \epsilon\right) < 1 \text{ if } 0 \notin [b_1, b_2]. \tag{5.67}$$

and

$$P\left(\log\left|1 + B \cdot k\right| < -\epsilon\right) < 1 \text{ if } 0 \in [b_1, b_2]. \tag{5.68}$$

The proof is in Appendix 5.8.

Lemma 5.4.7. *Suppose the system state at time n is $x[n] \in \mathbb{R}, x[n] \neq 0$. Then, for the system (5.46) define*

$$C_{ze,n}(x[n]) = \max\left\{C \middle| \max_{U[n](x[n])} P\left(-\log\left|\frac{X[n+1]}{x[n]}\right| > C\right) = 1\right\}. \tag{5.69}$$

Here, the probability is taken over the random realization of $B[n] \sim p_B(\cdot)$ i.i.d. with essential support $[b_1, b_2]$, and b_1 as the essential infimum and b_2 as the essential supremum of the *distribution of the random variables $B[n]$. Hence, $X[n+1]$ is a random variable even though $X[n] = x[n]$ has been realized and fixed. Then, $C_{ze,n}(x[n])$ does not depend on $x[n]$, and is given by*

$$C_{ze,n}(x[n]) = \begin{cases} \log\left|\frac{b_2+b_1}{b_2-b_1}\right| \textit{ if } \ 0 \notin [b_1, b_2], \ \textit{i.e. } 0 < b_1 < b_2 \textit{ or } b_1 < b_2 < 0 \\ 0 \textit{ if } 0 \in [b_1, b_2] \textit{ i.e. } b_1 \leq 0 \textit{ and } b_2 > 0, \ \textit{or } b_1 < 0 \textit{ and } b_2 \geq 0. \end{cases} \tag{5.70}$$

The achievability part of this lemma implies that there exists a linear control law $U[n]$,

$$P\left(-\log\left|\frac{X[n+1]}{x[n]}\right| \geq C_{ze,n}\right) = 1. \tag{5.71}$$

The converse implies that for any $\delta > 0$, for all $U[n]$,

$$P\left(-\log\left|\frac{X[n+1]}{x[n]}\right| > C_{ze,n} + \delta\right) < 1. \tag{5.72}$$

Proof: Achievability:
The case where $X[n+1] = 0$ is trivial, so we focus on the other case. Also, if $0 \in [b_1, b_2]$, choose $U[n] = 0$, which gives the result.

For the case $0 \notin [b_1, b_2]$, we would like to show that $P\left(-\log\left|\frac{X[n+1]}{x[n]}\right| \geq \log\left|\frac{b_1+b_2}{b_2-b_1}\right|\right) = 1$. Consider the following strategy for any $x[n] \neq 0$, choose $U[n] = -\left(\frac{b_1+b_2}{2}\right)^{-1} x[n]$. Then,

$$X[n+1] = x[n] - B[n]\left(\frac{b_1+b_2}{2}\right)^{-1} x[n] \tag{5.73}$$

$$= x[n]\left(1 - \frac{2B[n]}{b_1+b_2}\right). \tag{5.74}$$

Now, note that the following inequalities are equivalent.

$$-\log\left|\frac{X[n+1]}{x[n]}\right| \geq \log\left|\frac{b_1+b_2}{b_2-b_2}\right| \tag{5.75}$$

$$\Longleftrightarrow \log\left|\frac{X[n+1]}{x[n]}\right| \leq \log\left|\frac{b_2-b_1}{b_1+b_2}\right| \tag{5.76}$$

$$\Longleftrightarrow \left|1 - \frac{2B[n]}{b_1+b_2}\right| \leq \left|\frac{b_2-b_1}{b_1+b_2}\right| \tag{5.77}$$

Hence, the event $\mathcal{E}_1 = \left\{-\log\left|\frac{X[n+1]}{x[n]}\right| \geq \log\left|\frac{b_1+b_2}{b_2-b_1}\right|\right\}$, is identical to the event $\mathcal{E}_2 = \left\{\left|1 - \frac{2B[n]}{b_1+b_2}\right| \leq \left|\frac{b_2-b_1}{b_1+b_2}\right|\right\}$.

But we know from Lemma 5.4.6 that the maximum of $\left|1 - \frac{2B[n]}{b_1+b_2}\right|$ is $\left|\frac{b_2-b_1}{b_1+b_2}\right|$.

Hence, the event $\mathcal{E}_2$ occurs with probability 1, which implies $\mathcal{E}_1$ occurs with probability 1, which proves the achievability.

Converse: Here, we must show $\forall x[n] \neq 0, \forall \epsilon > 0$, for any $U[n](x[n])$,

$$P\left(-\log\left|\frac{X[n+1]}{x[n]}\right| > \log\left|\frac{b_1+b_2}{b_2-b_1}\right| + \epsilon\right) < 1. \tag{5.78}$$

- Case 1) $0 < b_1 < b_2$ or $b_1 < b_2 < 0$

$$\begin{aligned}
&P\left(-\log\left|\frac{X[n+1]}{x[n]}\right| > \log\left|\frac{b_2+b_1}{b_2-b_1}\right| + \epsilon\right) && (5.79)\\
=&P\left(\log\left|\frac{X[n+1]}{x[n]}\right| < \log\left|\frac{b_2-b_1}{b_1+b_2}\right| - \epsilon\right) && (5.80)\\
=&P\left(\log\left|\frac{x[n]+B[n]U[n]}{x[n]}\right| < \log\left|\frac{b_2-b_1}{b_1+b_2}\right| - \epsilon\right) && (5.81)\\
=&P\left(\log\left|1+B[n]\cdot k\right| < \log\left|\frac{b_2-b_1}{b_1+b_2}\right| - \epsilon\right) \ \text{(Let } \frac{U[n]}{x[n]} = k, \text{ for any strategy } U[n]) && (5.82)\\
<& 1 \text{ (by Lemma 5.4.6)} && (5.83)
\end{aligned}$$

- Case 2) $b_1 \le 0 < b_2$ or $b_1 < 0 \le b_2$

$$\begin{aligned}
&P\left(-\log\left|\frac{X[n+1]}{x[n]}\right| > \epsilon\right) && (5.84)\\
=&P\left(\log\left|\frac{X[n+1]}{X[n]}\right| < -\epsilon\right) && (5.85)\\
=&P\left(\log\left|\frac{X[n]+B[n]U[n]}{X[n]}\right| < -\epsilon\right) && (5.86)\\
=&P\left(\log\left|1+B[n]\cdot k\right| < -\epsilon\right) \quad \text{(Let } \frac{U[n]}{x[n]} = k, \text{ for any strategy } U[n]) && (5.87)\\
<& 1 \text{ (by Lemma 5.4.6)} && (5.88)
\end{aligned}$$

■

Remark 5.4.8. Lemma 5.4.7 can be extended to allow for randomized strategies $U[n]$.

Theorem 5.4.9. *Consider the system $\mathcal{S}(p_B)$ (parameterized with a single distribution $p_B(\cdot)$ since the $B[n]$'s are i.i.d.) given by eq. (5.46), with i.i.d $B[n] \sim p_B(\cdot)$, and where $p_B(\cdot)$*

is defined with essential support $[b_1, b_2]$*.* b_1 *is the essential infimum and* b_2 *is the essential supremum of the distribution of the random variables* $B[n]$*. Then the zero-error control capacity of the system is*

$$C_{ze}(\mathcal{S}) = \begin{cases} \log\left|\frac{b_2-b_1}{b_2+b_1}\right| & \text{if } 0 \notin [b_1, b_2] \\ 0 & \text{if } 0 \in [b_1, b_2] \end{cases} \tag{5.89}$$

Proof:
Achievability: Lemma 5.4.7 provides a sequence of $U[n]$ *for every* $x[n]$. If at any n we set the system to 0, we are clearly done. Otherwise, for every n, $P\left(-\log\left|\frac{X[n+1]}{X[n]}\right| \geq C_{ze}\right) = 1$. We apply these $U[n]$. Of course, if any $X[i] = 0$ for $1 \leq i \leq n$ we are done, since $X[i] = 0$ means we can have $X[i+1] = 0$ and we can define $\frac{0}{0} = 1$ to give us

Then:

$$P\left(-\frac{1}{n}\log\left|\frac{X[n]}{X[0]}\right| \geq C_{ze}\right) \tag{5.90}$$

$$=P\left(-\frac{1}{n}\sum_{i=0}^{n-1}\log\left|\frac{X[i+1]}{X[i]}\right| \geq C_{ze}\right) \tag{5.91}$$

$$=P\left(\sum_{i=0}^{n-1} -\log\left|\frac{X[i+1]}{X[i]}\right| \geq n \cdot C_{ze}\right) \tag{5.92}$$

$$\geq P\left(\left\{-\log\left|\frac{X[1]}{X[0]}\right| \geq C_{ze}\right\}, \cdots, \left\{-\log\left|\frac{X[n]}{X[n-1]}\right| > C_{ze}\right\}\right) \tag{5.93}$$

(The probability of the intersection of the events $\bigcap_{i=0}^{n-1}\left\{-\log\left|\frac{X[i+1]}{X[i]}\right| > C_{ze}\right\}$ is a lower bound on the desired probability.)

$$=1 \quad \text{(with the appropriate choice of } U[n] \text{ as above, apply Lemma 5.4.7.)} \tag{5.94}$$

Converse: To prove the converse, we would like to show $\forall\delta > 0$:

$$P\left(-\frac{1}{n}\log\left|\frac{X[n]}{X[0]}\right| > C_{ze} + \delta\right) < 1. \tag{5.95}$$

This is equivalent to showing that

$$P\left(-\frac{1}{n}\log\left|\frac{X[n]}{X[0]}\right| \leq C_{ze} + \delta\right) > 0. \tag{5.96}$$

We will use induction to show this. $n = 1$ is the base case. We know from Lemma 5.4.7 that

$$P\left(-\log\left|\frac{X[1]}{X[0]}\right| \leq C_{ze} + \delta\right) > 0. \tag{5.97}$$

Now, consider

$$
\begin{aligned}
&P\left(-\frac{1}{n}\log\left|\frac{X[n]}{X[0]}\right| \le C_{ze}+\delta\right) && (5.98)\\
=&P\left(-\log\left|\frac{X[n]}{X[0]}\right| \le n(C_{ze}+\delta)\right) && (5.99)\\
=&P\left(\sum_{i=0}^{n-1}-\log\left|\frac{X[i+1]}{X[i]}\right| \le n(C_{ze}+\delta)\right) && (5.100)
\end{aligned}
$$

Here, if $X[i]=0$ for any i, the best control is $U[i]=0$, and hence $X[n+1]=0$. We define $\frac{0}{0}=1$ in this context, and we are allowed to divide by $X[i]$ to get the expression above. We explicitly account for the probability that $X[i]$ might be zero later.

Now, the probability of this event is lower bounded by the joint probability of the two events below.

$$
\begin{aligned}
&\ge P\left(\left\{\textstyle\sum_{i=0}^{n-2}-\log\left|\frac{X[i+1]}{X[i]}\right| \le (n-1)(C_{ze}+\delta)\right\},\left\{-\log\left|\frac{X[n]}{X[n-1]}\right| \le (C_{ze}+\delta)\right\}\right) && (5.101)\\
&=\int P\left(\left\{\textstyle\sum_{i=0}^{n-2}-\log\left|\frac{X[i+1]}{X[i]}\right| \le (n-1)(C_{ze}+\delta)\right\},\left\{-\log\left|\frac{X[n]}{X[n-1]}\right| \le (C_{ze}+\delta)\right\}\middle|X[n-1]=q\right)\cdot p_{X[n-1]}(q)\ dq && (5.102)\\
&=\int\Bigg[P\left(\sum_{i=0}^{n-2}-\log\left|\frac{X[i+1]}{X[i]}\right| \le (n-1)(C_{ze}+\delta)\middle|X[n-1]=q\right)\cdot\\
&\qquad P\left(-\log\left|\frac{X[n]}{X[n-1]}\right| \le (C_{ze}+\delta)\middle|X[n-1]=q\right)\cdot p_{X[n-1]}(q)\Bigg]\ dq && (5.103)
\end{aligned}
$$

Conditioned on $X[n-1]$ the two events are independent.

Now, we can pull out the case of $X[n-1]=0$.

$$= \int_{X[n-1]\neq 0} \left[P\left(\sum_{i=0}^{n-2} -\log\left|\frac{X[i+1]}{X[i]}\right| \leq (n-1)(C_{ze}+\delta) \middle| X[n-1]=q \right) \cdot \right.$$
$$\left. P\left(-\log\left|\frac{X[n]}{X[n-1]}\right| \leq (C_{ze}+\delta) \middle| X[n-1]=q \right) \cdot p_{X[n-1]}(q) \right] \, dq$$
$$+P(X[n-1]=0)\cdot \left(P\left(\sum_{i=0}^{n-2} -\log\left|\frac{X[i+1]}{X[i]}\right| \leq (n-1)(C_{ze}+\delta) \middle| X[n-1]=0 \right) \cdot P\left(-\log\left|\frac{X[n]}{X[n-1]}\right| \leq (C_{ze}+\delta) \middle| X[n-1]=0 \right) \right) \tag{5.104}$$

$$= \int_{X[n-1]\neq 0} \left[P\left(\sum_{i=0}^{n-2} -\log\left|\frac{X[i+1]}{X[i]}\right| \leq (n-1)(C_{ze}+\delta) \middle| X[n-1]=q \right) \cdot \right.$$
$$\left. P\left(-\log\left|\frac{X[n]}{X[n-1]}\right| \leq (C_{ze}+\delta) \middle| X[n-1]=q \right) \cdot p_{X[n-1]}(q) \right] \, dq \tag{5.105}$$

(since $P\left(\sum_{i=0}^{n-2} -\log\left|\frac{X[i+1]}{X[i]}\right| \leq (n-1)(C_{ze}+\delta) \middle| X[n-1]=0 \right) = P(\infty \leq (n-1)(C_{ze}+\delta)) = 0$)

$$= \int_{X[n-1]\neq 0} \left[P\left(\sum_{i=0}^{n-2} -\log\left|\frac{X[i+1]}{X[i]}\right| \leq (n-1)(C_{ze}+\delta) \middle| X[n-1]=q \right) \cdot \right.$$
$$\left. P\left(-\log\left|\frac{X[n]}{q}\right| \leq (C_{ze}+\delta) \right) \cdot p_{X[n-1]}(q) \right] \, dq$$
$$= \int_{X[n-1]\neq 0} \left[P\left(\sum_{i=0}^{n-2} -\log\left|\frac{X[i+1]}{X[i]}\right| \leq (n-1)(C_{ze}+\delta), \{X[n-1]=q\} \right) \cdot \right.$$
$$\left. P\left(-\log\left|\frac{X[n]}{q}\right| \leq (C_{ze}+\delta) \right) \ \right] \, dq \tag{5.106}$$

Now, Lemma 5.4.7 implies that $P\left(-\log\left|\frac{X[n]}{q}\right| \leq (C_{ze}+\delta) = \right) > 0$ for all $q \neq 0$. The induction hypothesis implies that

$$\int_{X[n-1]\neq 0} P\left(\sum_{i=0}^{n-2} -\log\left|\frac{X[i+1]}{X[i]}\right| \leq (n-1)(C_{ze}+\delta) \middle| X[n-1]=q \right) dq \tag{5.107}$$
$$= P\left(\sum_{i=0}^{n-2} -\log\left|\frac{X[i+1]}{X[i]}\right| \leq (n-1)(C_{ze}+\delta) \right) \tag{5.108}$$
$$> 0. \tag{5.109}$$

Now, for non-negative functions f, g such that $f(q) \geq 0$, $\int f(q)dq > 0$ and $g(q) > 0$ for all q, we know that $\int f(q)g(q)dq > 0$. Applying this, we get,

$$P\left(\sum_{i=0}^{n-1} -\log\left|\frac{X[i+1]}{X[i]}\right| \leq n(C_{ze}+\delta)\right) > 0, \tag{5.110}$$

which proves the result. ■

Remark 5.4.10. This proof reveals that there is nothing lost by going to linear strategies for the zero-error control capacity problem. The core reason is the key lemma which lets us decouple the impact of the control at each time step. Furthermore, the fact that linear strategies suffice tells us that the zero-error control capacity result immediately generalizes to the case with bounded additive disturbances W and bounded additive noise V. The converse clearly holds because the additive terms can be given by a genie. The achievability also holds because the system in closed-loop becomes linear time-varying (where the time-variation is random). Because the zero-error capacity is strictly greater than $\log(a)$, the closed-loop dynamics will take all the initial conditions to zero exponentially fast. Where there are additive disturbances and noises, the resulting system state is a superposition of the responses to bounded inputs. The system is thus BIBO stable.

5.5 "Shannon" control capacity

Here we consider the stability of the logarithm of the state. Recall the definitions from earlier:

Definition 5.2.2. Let the evolution of system state $X[n]$ be governed by a linear system

$$\begin{aligned} X[n+1] &= A[n]\cdot X[n] + B[n]\cdot U[n] + W[n] \\ Y[n] &= C[n]\cdot X[n] + V[n] \end{aligned} \tag{5.4}$$

$A[n], B[n], C[n], V[n], W[n]$ are random variables with known distributions at each time n. Let $\sigma(Y_0^n)$ be the sigma-algebra generated by the random variables $Y[0]$ to $Y[n]$. The system is said to be **logarithmically stabilizable** if there exists a causal control strategy $U_0^n(\cdot) \in \sigma(Y_0^n)$ such that $\exists M \in \mathbb{R}$, $M < \infty$, s.t. $\mathbb{E}[\log|X[n]|] < M$ for all n.

Definition 5.5.1. The control capacity in the Shannon-sense of the system $\mathcal{S}(p_{B[0]}, p_{B[1]}, \cdots)$ in eq. (5.46) is defined as

$$C_{sh}(\mathcal{S}) = \lim_{n\to\infty} \; -\min_{U_0^{n-1}(\cdot)\ \text{s.t.}\ U[i](\cdot)\in\sigma(Y_0^n)} \frac{1}{n}\mathbb{E}\left[\log\left|\frac{X[n]}{X[0]}\right|\right]. \tag{5.111}$$

We also call this the zeroth-moment control capacity of the system, since it essentially captures the weakest notion of system stability.

Theorem 5.5.2. *Consider the system $\widetilde{\mathcal{S}}^a(p_B)$ (eq. (5.48)) and the affiliated system $\mathcal{S}(p_B)$ (eq. (5.46)), with identical initial conditions $X[0] = \widetilde{X}[0] \neq 0$, such that $-\infty < \mathbb{E}[\log |X[0]|] < \infty$, and identical $B[i]$'s.*

Then $\exists \widetilde{U}_0^\infty(\cdot)$, such that $\widetilde{U}[i](\cdot) \in \sigma(\widetilde{Y}_0^i)$ such that $\exists M, N < \infty$, $\mathbb{E}[\log |\widetilde{X}[n]|] < M$ for all $n > N$, i.e the system is logarithmically stabilizable if and only if $C_{sh}(\mathcal{S}(p_B)) \geq \log |a|$.

Proof: Recall the Lemma 5.4.4, from the zero-error capacity section, which says that the systems $\widetilde{\mathcal{S}}^a$ and $\widetilde{S}$ can be made to track each other. This lemma will be used here as well.

($\Longrightarrow$) If we know there exists $\widetilde{U}_0^\infty(\cdot)$ such that $\mathbb{E}\left[\log |\widetilde{X}[n]|\right] < M, \forall\ n > N$.

$$\mathbb{E}\left[\log |\widetilde{X}[n]|\right] < M, \forall\ n \tag{5.112}$$

$$\Longrightarrow \mathbb{E}\left[\log \frac{|\widetilde{X}[n]|}{|X[0]|}\right] < M - \mathbb{E}\left[\log |X[0]|\right], \forall\ n > N \text{ (since } \mathbb{E}[\log |X[0]|] < \infty) \tag{5.113}$$

$$\Longrightarrow \mathbb{E}\left[\log \frac{|a|^n |X[n]|}{|X[0]|}\right] < M - \mathbb{E}\left[\log |X[0]|\right], \forall\ n > N \tag{5.114}$$

(Choose $U[i] = a^{-i}\widetilde{U}[i]$ and apply Lemma 5.4.4)

$$\Longrightarrow n \log |a| + \mathbb{E}\left[\log \frac{|X[n]|}{|X[0]|}\right] < M - \mathbb{E}\left[\log |X[0]|\right], \forall\ n > N \tag{5.115}$$

$$\Longrightarrow \log |a| + \frac{1}{n}\mathbb{E}\left[\log \frac{|X[n]|}{|X[0]|}\right] < \frac{M - \mathbb{E}\left[\log |X[0]|\right]}{n}, \forall\ n > N \tag{5.116}$$

$$\Longrightarrow -\frac{1}{n}\mathbb{E}\left[\log \frac{|X[n]|}{|X[0]|}\right] > \log |a| - \frac{M - \mathbb{E}\left[\log |X[0]|\right]}{n}, \forall\ n > N \tag{5.117}$$

$$\Longrightarrow \liminf_{n\to\infty} -\frac{1}{n}\mathbb{E}\left[\log \frac{|X[n]|}{|X[0]|}\right] \geq \log |a| \quad \text{(the term } \mathbb{E}\left[\log |X[0]|\right] \text{ is bounded.)} \tag{5.118}$$

$$\Longrightarrow C_{sh}\left(\widetilde{S}(p_B)\right) \geq \log |a| \text{ (the optimal control must be at least as good as the applied control)} \tag{5.119}$$

($\Longleftarrow$)

$C_{sh}\left(S(p_B)\right) \geq \log |a|$ and so $\exists U_0^\infty(\cdot)$ such that $U[i](\cdot) \in \sigma(Y_0^i)$ and $N \in \mathbb{R}$ such that $\forall\ n > N$, $-\frac{1}{n}\mathbb{E}\left[\log \frac{|X[n]|}{|\widetilde{X}[0]|}\right] \geq \log |a|$. Now, choose $\widetilde{U}[i] = a^i U[i]$.

$$-\frac{1}{n}\mathbb{E}\left[\log\frac{|X[n]|}{|X[0]|}\right] \geq \log|a|, \forall\ n > N \tag{5.120}$$

$$\Longrightarrow \mathbb{E}\left[\log\frac{|X[n]|}{|X[0]|}\right] \leq -n\log|a|, \forall\ n > N \tag{5.121}$$

$$\Longrightarrow \mathbb{E}\left[\log\frac{|a|^n|X[n]|}{|X[0]|}\right] \leq 0, \forall\ n > N \tag{5.122}$$

$$\Longrightarrow \mathbb{E}\left[\log\frac{|\widetilde{X}[n]|}{|\widetilde{X}[0]|}\right] \leq 0, \forall\ n > N \tag{5.123}$$

$$\Longrightarrow \mathbb{E}\left[\log|\widetilde{X}[n]|\right] \leq \mathbb{E}\left[\log|\widetilde{X}[0]|\right] < \infty, \forall\ n > N \text{ (by assumption).} \tag{5.124}$$

■

5.5.1 Calculating Shannon control capacity

Theorem 5.5.3. *The Shannon control capacity of the system $\mathcal{S}(p_B)$ (eq. (5.46)), with i.i.d. $B[n]$'s, is ∞ if $p_B(\cdot)$ has an atom not at 0.*

Proof: Let $p_B(\cdot)$ have an atom at $\beta \neq 0$. Then consider the strategy $U[i] = -\frac{1}{\beta}X[i]$. In this case, $\frac{X[n]}{X[0]}$ can be 0 with positive probability, and hence the negative log can be infinite. So the control capacity is infinite. Essentially, this captures the idea that betting on the atom will eventually pay off if we wait long enough. ■

Theorem 5.5.4. *The Shannon control capacity of the system $\mathcal{S}(p_B)$ (eq. (5.46)), where $p_B(\cdot)$ is a continuous distribution with no atoms is given by,*

$$\max_{k\in\mathbb{R}} \mathbb{E}\left[-\log|1 + B\cdot k|\right] \tag{5.125}$$

where $B \sim p_B(\cdot)$.

Lemma 5.5.5. *Suppose the system state at time n is $x[n] \in \mathbb{R}, x[n] \neq 0$. Then, for the system $\mathcal{S}$ in eq. (5.46) define*

$$C_{sh,n}(x[n]) = -\min_{U[n](\cdot)} \mathbb{E}\left[\log\left|\frac{X[n+1]}{x[n]}\right|\right]. \tag{5.126}$$

where $U[n](\cdot)$ is any function of $x[n]$. Here, the expectation is taken over the random realization of $B[n] \sim p_B(\cdot)$ i.i.d.. Hence, $X[n+1]$ is a random variable even though $X[n] = x[n]$ has been realized and fixed. Then, $C_{sh,n}(x[n])$ does not depend on $x[n]$ or n, and is given by

$$C_{sh,n}(x[n]) = \max_{k\in\mathbb{R}} \mathbb{E}\left[-\log|1 + B[0]\cdot k|\right], \tag{5.127}$$

where $B[0] \sim p_B(\cdot)$. Hence, there exists $U[n](x[n])$ such that

$$\mathbb{E}\left[\log\left|\frac{x[n]}{X[n+1]}\right|\right] = \max_k \mathbb{E}\left[-\log|1 + B[0]\cdot k|\right]. \tag{5.128}$$

Proof:

$$\begin{aligned}
C_n &= -\min_{U[n](x[n])} \mathbb{E}\left[\log\left|\frac{X[n+1]}{x[n]}\right|\right] && (5.129)\\
&= \max_{U[n](x[n])} -\mathbb{E}\left[\log\left|\frac{X[n+1]}{x[n]}\right|\right] && (5.130)\\
&= \max_{U[n](x[n])} -\mathbb{E}\left[\log\left|1 + B[n]\cdot\frac{U[n](x[n])}{x[n]}\right|\right] && (5.131)\\
&= \max_{k(x[n])} \mathbb{E}\left[-\log\left|1 + B[n]\cdot k(x[n])\right|\right] \text{ (since } x[n] \neq 0 \text{ and } U[n] \text{ is a function of } x[n].) && (5.132)\\
&= \max_k \mathbb{E}\left[-\log\left|1 + B[n]\cdot k\right|\right] \text{ (}k\text{ is just a constant)} && (5.133)
\end{aligned}$$

■

We have one last small lemma before the proof.

Lemma 5.5.6. *Given that $p_{X[0]}(\cdot)$ has no atom at 0 and $p_B(\cdot)$ has no atoms, $\forall n$ $X[n]$ cannot have an atom at 0.*

Proof: We use induction. $X[0]$ is the base case. Note, that in the event $U[n] = 0$ is applied, $X[n+1] = X[n]$ and so $p_{X[n+1]}$ cannot have an atom at 0 by the induction hypothesis.

If $U[n] \neq 0$ is applied:

$$\begin{aligned}
&P(X[n+1] = 0)\\
=&P(X[n] + B[n]\cdot U[n] = 0) && (5.134)\\
=&P(B[n] = -\frac{X[n]}{U[n]}) && (5.135)\\
&(U[n] \text{ is a function of } X[n] \text{ and hence the ratio can be thought of as a constant})\\
=&0 && (5.136)
\end{aligned}$$

since $p_B(\cdot)$ has no atoms. ■

Finally, we can prove Thm. 5.5.4!

Proof: (Thm. 5.5.4)

Achievability: We know from Lemma 5.5.5 that there exists $U[n](x[n])$ such that

$$-\mathbb{E}\left[\log\left|\frac{X[n+1]}{x[n]}\right|\right] = \max_{k\in\mathbb{R}} \mathbb{E}\left[-\log|1+B[n]\cdot k|\right] \tag{5.137}$$

for every $x[n]$. Thus, we have a control law $U[0](\cdot)$ that gives the appropriate control for each realization of the initial state. We can choose to apply this sequence of $U[i](X[i])$ for $0 \le i \le n$ based on the realization of $X[0]$. We are allowed to divide by $X[i]$ below due to the Lemma 5.5.6 above.
Hence,

$$\lim_{n\to\infty} \quad -\min_{U_0^{n-1}(\cdot)\text{s.t. } U[i](\cdot)\in\sigma(Y_0^i)} \frac{1}{n}\mathbb{E}\left[\log\left|\frac{X[n]}{X[0]}\right|\right] \tag{5.138}$$

$$= \lim_{n\to\infty} \quad -\min_{U_0^{n-1}(\cdot)\text{s.t. } U[i](\cdot)\in\sigma(Y_0^i)} \frac{1}{n}\mathbb{E}\left[\sum_{i=0}^{n-1}\log\left|\frac{X[i+1]}{X[i]}\right|\right] \tag{5.139}$$

$$= \lim_{n\to\infty} \quad -\min_{U_0^{n-1}(\cdot)\text{s.t. } U[i](\cdot)\in\sigma(Y_0^i)} \frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}\left[\log\left|\frac{X[i+1]}{X[i]}\right|\right] \text{ (Linearity of expectation)} \tag{5.140}$$

$$= \lim_{n\to\infty} \quad -\min_{U_0^{n-1}(\cdot)\text{s.t. } U[i](\cdot)\in\sigma(Y_0^i)} \frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}_{X[i]}\left[\mathbb{E}_{B[i]}\left[\log\left|\frac{X[i+1]}{X[i]}\right|\Bigg|X[i]\right]\right] \tag{5.141}$$

(Since the $B[i]$'s are independent the ith term only has $B[i]$ as the source of randomness)

$$\ge \lim_{n\to\infty} \quad -\min_{U_0^{n-1}(\cdot)\text{s.t. } U[i](\cdot)\in\sigma(Y_0^i)} \frac{1}{n}\sum_{i=0}^{n-1}\max_{X[i]\neq 0}\mathbb{E}_{B[i]}\left[\log\left|\frac{X[i+1]}{X[i]}\right|\right] \tag{5.142}$$

(The worst case $X[i]$ realization lowerbounds performance, regardless of the past actions.)
(Further, we restrict to $X[i]\neq 0$ since there is no atom at 0 in $p_{X[0]}(\cdot)$.)
(Since there are no atoms in $p_B(\cdot)$, the actions cannot create an atom at 0 for any $X[i]$.)

$$= \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} -\min_{U[i](\cdot)\in\sigma(Y_0^i)}\max_{X[i]\neq 0}\mathbb{E}_{B[i]}\left[\log\left|\frac{X[i+1]}{X[i]}\right|\right] \tag{5.143}$$

(Assuming the worst case state at each time decouples the effect of the $U[i]$s.)

$$= \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} -\min_{U[i](\cdot)\in\sigma(Y[i])}\max_{X[i]\neq 0}\mathbb{E}_{B[i]}\left[\log\left|\frac{X[i+1]}{X[i]}\right|\right] \tag{5.144}$$

(The system is Markov, and $Y[i]$ captures the information from the entire past)

$$= \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1}\max_{k}\mathbb{E}\left[-\log\Big|1+B[i]\cdot k\Big|\right] \text{ (by Lemma 5.5.5)} \tag{5.145}$$

$$= \max_{k\in\mathbb{R}}\mathbb{E}\left[-\log\Big|1+B[0]\cdot k\Big|\right] \text{ (since the } B[i] \text{ are i.i.d.)} \tag{5.146}$$

Converse: Recall that we are allowed to divide by $X[i]$ below due to the Lemma 5.5.6 above.

$$\lim_{n\to\infty} \quad -\min_{U_0^{n-1}(\cdot)\text{s.t. } U[i](\cdot)\in\sigma(Y_0^i)} \frac{1}{n}\mathbb{E}\left[\log\left|\frac{X[n]}{X[0]}\right|\right] \tag{5.147}$$

$$= \lim_{n\to\infty} \quad -\min_{U_0^{n-1}(\cdot)\text{s.t. } U[i](\cdot)\in\sigma(Y_0^i)} \frac{1}{n}\mathbb{E}\left[\log\left|\frac{X[n]}{X[n-1]}\right| + \log\left|\frac{X[n-1]}{X[0]}\right|\right] \tag{5.148}$$

$$\leq \lim_{n\to\infty} \quad -\min_{U_0^{n-1}(\cdot)\text{s.t. } U[i](\cdot)\in\sigma(Y_0^i)} \frac{1}{n}\mathbb{E}\left[\log\left|\frac{X[n]}{X[n-1]}\right|\right] + -\min_{U_0^{n-1}(\cdot)\text{s.t. } U[i](\cdot)\in\sigma(Y_0^i)} \frac{1}{n}\mathbb{E}\left[\log\left|\frac{X[n-1]}{X[0]}\right|\right] \tag{5.149}$$

$$= \lim_{n\to\infty} \quad -\min_{U_0^{n-1}(\cdot)\text{s.t. } U[i](\cdot)\in\sigma(Y_0^i)} \frac{1}{n}\mathbb{E}\left[\log\left|\frac{X[n]}{X[n-1]}\right|\right] + -\min_{U_0^{n-2}\text{s.t. } U[i](\cdot)\in\sigma(Y_0^i)} \frac{1}{n}\mathbb{E}\left[\log\left|\frac{X[n-1]}{X[0]}\right|\right] \tag{5.150}$$

(remove the irrelevant parameters in the minimization)

$$= \lim_{n\to\infty} \quad -\min_{U_0^{n-1}(\cdot)\text{s.t. } U[i](\cdot)\in\sigma(Y_0^i)} \frac{1}{n}\mathbb{E}_{X[n-1]}\left[\mathbb{E}\left[\log\left|\frac{X[n]}{X[n-1]}\right|\middle|X[n-1]\right]\right] + -\min_{U_0^{n-2}\text{s.t. } U[i](\cdot)\in\sigma(Y_0^i)} \frac{1}{n}\mathbb{E}\left[\log\left|\frac{X[n-1]}{X[0]}\right|\right] \tag{5.151}$$

$$\leq \lim_{n\to\infty} \quad -\min_{U_0^{n-1}(\cdot)\text{s.t. } U[i](\cdot)\in\sigma(Y_0^i)} \frac{1}{n}\min_{X[n-1]\neq 0} \mathbb{E}\left[\log\left|\frac{X[n]}{X[n-1]}\right|\right] + -\min_{U_0^{n-2}\text{s.t. } U[i](\cdot)\in\sigma(Y_0^i)} \frac{1}{n}\mathbb{E}\left[\log\left|\frac{X[n-1]}{X[0]}\right|\right] \tag{5.152}$$

(Since we consider the best case over $X[n-1]$, and $B[i]$'s are i.i.d., prior history is irrelevant.)

$$\leq \lim_{n\to\infty} \quad -\min_{U[n-1](\cdot)\in\sigma(Y_0^{n-1})} \frac{1}{n}\min_{X[n-1]\neq 0} \mathbb{E}\left[\log\left|\frac{X[n]}{X[n-1]}\right|\right] + -\min_{U_0^{n-2}\text{s.t. } U[i](\cdot)\in\sigma(Y_0^i)} \frac{1}{n}\mathbb{E}\left[\log\left|\frac{X[n-1]}{X[0]}\right|\right] \tag{5.153}$$

$$= \lim_{n\to\infty} \quad -\min_{U[n-1](\cdot)\in\sigma(Y[n-1])} \frac{1}{n}\min_{X[n-1]\neq 0} \mathbb{E}\left[\log\left|\frac{X[n]}{X[n-1]}\right|\right] + -\min_{U_0^{n-2}\text{s.t. } U[i](\cdot)\in\sigma(Y_0^i)} \frac{1}{n}\mathbb{E}\left[\log\left|\frac{X[n-1]}{X[0]}\right|\right] \tag{5.154}$$

(Since the system is Markov we can restrict $U[n-1](\cdot)\in\sigma(Y[n-1])$.)

$$= \lim_{n\to\infty} \frac{1}{n}\max_k \mathbb{E}\left[-\log\left|1+B[n-1]\cdot k\right|\right] + -\min_{U_0^{n-2}\text{s.t. } U[i](\cdot)\in\sigma(Y_0^i)} \frac{1}{n}\mathbb{E}\left[\log\left|\frac{X[n-1]}{X[0]}\right|\right] \tag{5.155}$$

(by Lemma 5.5.5)

$$= \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n}\max_k \mathbb{E}\left[-\log\left|1+B[i]\cdot k\right|\right] \quad \text{(by induction)} \tag{5.156}$$

$$= \max_{k\in\mathbb{R}} \mathbb{E}\left[-\log\left|1+B[0]\cdot k\right|\right] \tag{5.157}$$

■

Remark 5.5.7. Once again, as a consequence of the proof we see that linear strategies are optimal for calculating the Shannon control capacity. As in the zero-error case, this immediately gives us what we want even for the case of additive disturbances. This is because when the Shannon control-capacity is $> \log(a)$, the typical LTV sample paths realized by the closed-loop system are decaying exponentially and are hence BIBO stable.

Remark 5.5.8. To compute the Shannon control capacity of distributions $p_B(\cdot)$ with atoms at zero, we can remove the atom at zero and renormalize the result as $P(B \neq 0) \cdot C_{sh}\left(\mathcal{S}(p_{B^*}(\cdot))\right)$, where p_{B^*} is the renormalized distribution after removing the atom at 0.

5.6 η-th moment control capacity

Finally, we can define the η-th moment stability. As η ranges from 0 to ∞ it captures a range of stabilities from the weakest "Shannon" sense as $\eta \to 0$, to the zero-error notions of stability as $\eta \to \infty$.

Definition 5.6.1. Let the evolution of system state $X[n]$ be governed by a linear system

$$\begin{aligned} X[n+1] &= A[n] \cdot X[n] + B[n] \cdot U[n] + W[n] \\ Y[n] &= C[n] \cdot X[n] + V[n] \end{aligned} \tag{5.158}$$

$A[n], B[n], C[n], V[n], W[n]$ are random variables with known distributions. Let $\sigma(Y_0^n)$ be the sigma-algebra generated by the random variables $Y[0]$ to $Y[n]$. The system is said to be **η-th moment stabilizable** if there exists a causal control strategy $U_0^n(\cdot) \in \sigma(Y_0^n)$ such that $\exists M, N$ such that $M, N < \infty$ and $\mathbb{E}[|X[n]|^\eta] < M$ for all $n > N$.

Definition 5.6.2. The η-th moment control capacity of the system $\mathcal{S}(p_{B[0]}, p_{B[1]}, \cdots)$ as in eq. (5.46) is

$$C_\eta(\mathcal{S}) = \liminf_{n\to\infty} \quad - \min_{U_0^{n-1}(\cdot) \text{s.t.} U[i](\cdot) \in \sigma(Y_0^i)} \frac{1}{n} \frac{1}{\eta} \log \mathbb{E}\left[\left|\frac{X[n]}{X[0]}\right|^\eta\right] \tag{5.159}$$

Theorem 5.6.3. *Consider the system $\widetilde{\mathcal{S}}^a(p_{B[0]}, p_{B[1]}, \cdots)$ from eq. (5.48), with initial condition $\widetilde{X}[0]$ such that $p_{\widetilde{X}[0]}(\cdot)$ has bounded support, i.e. $\exists \gamma < \infty$ such that $|X[0]| < \gamma$, and has no atom at* 0.

1. *If the η-th moment control capacity of the associated system $\mathcal{S}(p_{B[0]}, p_{B[1]}, \cdots)$ as in eq. (5.46) is high enough, i.e. $C_\eta(\mathcal{S}) \geq \log|a|$, then $\widetilde{\mathcal{S}}^a(p_{B[0]}, p_{B[1]}, \cdots)$ is η-th moment stabilizable.*

2. *Further, if* $\widetilde{\mathcal{S}}^a(p_{B[0]}, p_{B[1]}, \cdots)$ *is* η*-th moment stabilizable for every initial condition* $X[0] = x_0$*, i.e.* $\exists M(x_0)$ *such that* $\mathbb{E}[|X[n]|^\eta | X[0] = x_0] < M(x_0)$*, and* $\mathbb{E}\left[\frac{M(x_0)}{|X[0]|^\eta}\right] < \infty$ *then* $C_\eta(\mathcal{S}) \geq \log |a|$.

This motivates the definition of η*-th moment control capacity for the system* $\widetilde{\mathcal{S}}^a(p_{B[0]}, p_{B[1]}, \cdots)$ *to be the same as that of the system* $\mathcal{S}(p_{B[0]}, p_{B[1]}, \cdots)$.

Before diving into the proof, we recall the lemma:

Lemma 5.4.4. *Let* U_0^∞ *be a set of controls applied to* $\mathcal{S}(p_{B[0]}, p_{B[1]}, \cdots)$*. Let* $\widetilde{U}[i] = a^i U[i]$ *be the controls applied to* $\widetilde{\mathcal{S}}^a(p_{B[0]}, p_{B[1]}, \cdots)$*, with the same realizations of the random variables* $B[i]$ *and initial condition* $X[0]$*. Then,* $\widetilde{U}[i](\cdot) \in \sigma(\widetilde{Y}_0^i)$ *and* $\forall\ i, \widetilde{X}[i] = a^i X[i]$.

Proof: (**Thm. 5.6.3**)

Proof of the first claim of Thm. 5.6.3: We know there exists a control strategy $U_0^\infty(\cdot)$ and $N < \infty$ such that for $n > N$

$$-\frac{1}{n}\frac{1}{\eta}\log \mathbb{E}\left[\left|\frac{X[n]}{X[0]}\right|^\eta\right] \geq |a| \tag{5.160}$$

$$\iff \mathbb{E}\left[\left|\frac{X[n]}{X[0]}\right|^\eta\right] \leq |a|^{-n\eta} \tag{5.161}$$

$$\iff \mathbb{E}\left[\left|\frac{\widetilde{X}[n]}{X[0]}\right|^\eta\right] \leq 1 \text{ (Using Lemma 5.4.4, and linearity of expectation).} \tag{5.162}$$

$$\iff \mathbb{E}_{X[0]}\left[\frac{1}{|X[0]|^\eta}\mathbb{E}\left[|\widetilde{X}[n]|^\eta \middle| X[0]\right]\right] \leq 1 \tag{5.163}$$

$$\implies \mathbb{E}_{X[0]}\left[\frac{1}{|\gamma|^\eta}\mathbb{E}\left[|\widetilde{X}[n]|^\eta \middle| X[0]\right]\right] \leq 1 \text{ (since } |X[0]| < \gamma) \tag{5.164}$$

$$\implies \mathbb{E}_{X[0]}\left[\mathbb{E}\left[|\widetilde{X}[n]|^\eta \middle| X[0]\right]\right] \leq \gamma^\eta \tag{5.165}$$

$$\implies \mathbb{E}\left[|\widetilde{X}[n]|\right] \leq \gamma^\eta \tag{5.166}$$

which gives the required bound M. ■

Proof of the second claim of Thm. 5.6.3: We know there exists a control strategy $\widetilde{U}_0^\infty(\cdot)$ and an $M(x_))$ function for all x_0, such that for $\forall n > N$

$$\mathbb{E}\left[|\widetilde{X}[n]|\middle|X[0]=x_0\right] \leq M(x_0) \tag{5.167}$$

$$\Longrightarrow \mathbb{E}\left[\frac{|\widetilde{X}[n]|^\eta}{|x_0|^\eta}\middle|X[0]=x_0\right] \leq \frac{M(x_0)}{|x_0|^\eta} \tag{5.168}$$

$$\Longrightarrow \mathbb{E}_{X[0]}\left[\mathbb{E}\left[\frac{|\widetilde{X}[n]|^\eta}{|X[0]|^\eta}\middle|X[0]\right]\right] \leq \mathbb{E}_{X[0]}\left[\frac{M(X[0])}{|X[0]|^\eta}\right] \tag{5.169}$$

$$\Longrightarrow \mathbb{E}\left[\frac{|\widetilde{X}[n]|^\eta}{|X[0]|^\eta}\right] \leq \mathbb{E}\left[\frac{M(X[0])}{|X[0]|^\eta}\right] \tag{5.170}$$

$$\Longrightarrow |a|^{n\eta}\mathbb{E}\left[\frac{|X[n]|^\eta}{|X[0]|^\eta}\right] \leq \mathbb{E}\left[\frac{M(X[0])}{|X[0]|^\eta}\right] \text{ (Using Lemma 5.4.4)} \tag{5.171}$$

$$\Longrightarrow \mathbb{E}\left[\left|\frac{X[n]}{X[0]}\right|^\eta\right] \leq |a|^{-n\eta}\mathbb{E}\left[\frac{M(X[0])}{|X[0]|^\eta}\right] \tag{5.172}$$

$$\Longrightarrow \log\mathbb{E}\left[\left|\frac{X[n]}{X[0]}\right|^\eta\right] \leq -n\eta\log|a| + \log\mathbb{E}\left[\frac{M(X[0])}{|X[0]|^\eta}\right] \tag{5.173}$$

$$\Longrightarrow -\frac{1}{n}\frac{1}{\eta}\log\mathbb{E}\left[\left|\frac{X[n]}{X[0]}\right|^\eta\right] \geq \log|a| - \frac{1}{n}\frac{1}{\eta}\log\mathbb{E}\left[\frac{M(X[0])}{|X[0]|^\eta}\right] \tag{5.174}$$

$$\Longrightarrow \liminf_{n\to\infty} -\frac{1}{n}\frac{1}{\eta}\log\mathbb{E}\left[\left|\frac{X[n]}{X[0]}\right|^\eta\right] \geq \log|a| \tag{5.175}$$

$$\Longrightarrow C_\eta(\mathcal{S}) \geq \log|a| \tag{5.176}$$

■

5.6.1 Calculating η-th moment control capacity

This calculation uses a key one-step lemma, just as in the Shannon and zero-error cases.

Lemma 5.6.4. *Suppose the system state at time n is $x[n] \in \mathbb{R}, x[n] \neq 0$. Then, for system $\mathcal{S}(p_B)$ in eq. (5.46) define*

$$C_{\eta,n}(x[n]) = -\min_{U[n](\cdot)} \frac{1}{\eta}\log\mathbb{E}\left[\left|\frac{X[n+1]}{x[n]}\right|^\eta\right]. \tag{5.177}$$

where $U[n](\cdot)$ is any function of $x[n]$. Here, the expectation is taken over the random realization of $B[n] \sim p_B(\cdot)$ i.i.d.. Hence, $X[n+1]$ is a random variable even though $X[n] = x[n]$ has been realized and fixed. $Y[n] = y[n]$ is also fixed. Then, $C_{\eta,n}(x[n])$ does not depend on $x[n]$ or n, and is given by

$$C_{\eta,n}(x[n]) = -\min_{k\in\mathbb{R}} \frac{1}{\eta}\log\mathbb{E}\left[|1 + B[n]\cdot k|^\eta\right], \tag{5.178}$$

where $B[n] \sim p_B(\cdot)$. Hence, there exists $U[n](x[n])$ such that

$$-\log \mathbb{E}\left[\left|\frac{X[n+1]}{x[n]}\right|\right] = -\min_{k\in\mathbb{R}} \log \mathbb{E}\left[|1 + B[n]\cdot k|^{\eta}\right]. \tag{5.179}$$

Proof: This proof is essentially identical to the proof of Lemma 5.5.5 in the Shannon case. ■

Theorem 5.6.5. *The η-th moment control capacity of the system $\mathcal{S}(p_B)$, parameterized with a single distribution $p_B(\cdot)$ with $B[n]$s are i.i.d. , and with $p_B(\cdot)$ a continuous distribution with no atoms is given by*

$$C_\eta = -\min_{k\in\mathbb{R}} \frac{1}{\eta}\log \mathbb{E}\left[|1 + B\cdot k|^{\eta}\right] \tag{5.180}$$

where $B \sim p_B(\cdot)$.

Proof: The achievability is very similar to that of the Shannon case. We use the linear law from Lemma 5.6.4 and take advantage of the independence of the $B[i]$'s to turn an expectation of a product into a product of expectations.
Converse: Consider

$$\min_{U_0^{n-1}(\cdot)\text{ s.t. } U[i](\cdot)\in\sigma(Y_0^i)} - \mathbb{E}\left[\left|\frac{X[n]}{X[0]}\right|^\eta\right] \tag{5.181}$$

$$= \min_{U_0^{n-1}(\cdot)\text{ s.t. } U[i](\cdot)\in\sigma(Y_0^i)} - \mathbb{E}\left[\left|\frac{X[n]}{X[n-1]}\right|^\eta\left|\frac{X[n-1]}{X[0]}\right|^\eta\right] \tag{5.182}$$

(Since $p_B(\cdot)$ has no atoms, using Lemma 5.5.6, $X[n-1]$ cannot have an atom at 0)

$$= \min_{U_0^{n-1}(\cdot)\text{ s.t. } U[i](\cdot)\in\sigma(Y_0^i)} - \mathbb{E}_{X[0]}\mathbb{E}_{B_0^{n-2}}\mathbb{E}_{B[n-1]}\left[\left|\frac{X[n]}{X[n-1]}\right|^\eta\left|\frac{X[n-1]}{X[0]}\right|^\eta\middle|X[0],B_0^{n-2}\right] \tag{5.183}$$

$$= \min_{U_0^{n-1}(\cdot)\text{ s.t. } U[i](\cdot)\in\sigma(Y_0^i)} - \mathbb{E}_{X[0]}\mathbb{E}_{B_0^{n-2}}\left[\left|\frac{X[n-1]}{X[0]}\right|^\eta\mathbb{E}_{B[n-1]}\left[\left|\frac{X[n]}{X[n-1]}\right|^\eta\middle|X[0],B_0^{n-2}\right]\right] \tag{5.184}$$

$$\leq \min_{U_0^{n-1}(\cdot)\text{ s.t. } U[i](\cdot)\in\sigma(Y_0^i)} - \mathbb{E}_{X[0]}\mathbb{E}_{B_0^{n-2}}\left[\left|\frac{X[n-1]}{X[0]}\right|^\eta\mathbb{E}_{B[n-1]}\left[\min_{X[n-1]\neq 0}\left|\frac{X[n]}{X[n-1]}\right|^\eta\middle|X[0],B_0^{n-2}\right]\right] \tag{5.185}$$

(The best possible $X[n-1]$ as a function of $X[0]$ and history gives a lower bound.)

$$= \min_{U_0^{n-2}(\cdot)\text{ s.t. } U[i](\cdot)\in\sigma(Y_0^i)} - \mathbb{E}_{X[0]}\mathbb{E}_{B_0^{n-2}}\left[\left|\frac{X[n-1]}{X[0]}\right|^\eta\mathbb{E}_{B[n-1]}\left[\min_{U[n-1]\in\sigma(Y[n-1])}\min_{X[n-1]\neq 0}\left|\frac{X[n]}{X[n-1]}\right|^\eta\middle|X[0],B_0^{n-2}\right]\right]$$

$$= \min_{U_0^{n-2}(\cdot)\text{ s.t. } U[i](\cdot)\in\sigma(Y_0^i)} - \min_k \mathbb{E}\left[|1+B[n]\cdot k|^\eta\right]\cdot\mathbb{E}_{X[0]}\mathbb{E}_{B_0^{n-2}}\left[\left|\frac{X[n-1]}{X[0]}\right|^\eta\right] \tag{5.186}$$

$$= \left(-\min_{k\in\mathbb{R}}\mathbb{E}\left[|1+B[0]\cdot k|^\eta\right]\right)^n \tag{5.187}$$

(by induction, and the fact that the $B[i]$'s are i.i.d.)

So we have:

$$\lim_{n\to\infty}\min_{U_0^{n-1}(\cdot)\text{ s.t. } U[i](\cdot)\in\sigma(Y_0^i)} - \frac{1}{n}\frac{1}{\eta}\log\mathbb{E}\left[\left|\frac{X[n]}{X[0]}\right|^\eta\right] \tag{5.188}$$

$$\leq \lim_{n\to\infty}\min_{U_0^{n-1}(\cdot)\text{ s.t. } U[i](\cdot)\in\sigma(Y_0^i)} - \frac{1}{\eta}\log\max_k\mathbb{E}\left[-|1+B[0]\cdot k|^\eta\right] \tag{5.189}$$

$$= -\frac{1}{\eta}\log\min_{k\in\mathbb{R}}\mathbb{E}\left[|1+B[0]\cdot k|^\eta\right] \tag{5.190}$$

■

Again, as a side product of this result, we can see that there is no loss of optimality in restricting to linear strategies for the purposes of calculating η-th moment control capacity.

Remark 5.6.6. The proof uses the fact that there are no atoms, to avoid division by zero. However, a more careful treatment of the probability with which the system state becomes zero can be used to extend these ideas to cover the case of atoms as well.

Corollary 5.6.7. *Consider the system $\mathcal{S}(p_B)$ from (eq. (5.46)), with $B[n] \sim p_B(\cdot)$ i.i.d with mean μ_B and variance σ_B^2. Then*

$$C_2 = \frac{1}{2}\log\left(1 + \frac{\mu_B^2}{\sigma_B^2}\right). \tag{5.191}$$

Proof: We know that

$$C_2 = -\min_{k \in \mathbb{R}} \log \mathbb{E}\left[|1 + B \cdot k|^2\right] \tag{5.192}$$

Taking derivatives, the optimum $k = -\frac{\mu_B}{\mu_B^2 + \sigma_B^2}$, just as in the first example 5.2.5. Substituting k back into the equation gives the desired result. In the first example, we restricted our search space to linear strategies. Our converse tells us that linear strategies suffice. ■

Remark 5.6.8. This optimality of linear strategies for the second moment case was seen in the classical uncertainty threshold principle, but we can now say they are optimal from a control capacity perspective for all moments.

Remark 5.6.9. The similarity of this expression to the $\frac{1}{2}\log(1 + SNR)$ formula for communication capacity in the AWGN case is surprising. This is the notion of capacity that Elia hints at in [30]. Given the surprising coincidence in the value of second moment control capacity, the differences between communication capacity and control capacity were not noticed in the community. We believe that proof techniques here can be extended to understand stabilizability for a wider family of systems.

5.7 Examples

Here we calculate and plot the different notions of control capacity discussed in the chapter. Fig. 5.3 plots the zero-error, Shannon and second-moment control capacities for a Gaussian distribution, a Bernoulli-(p) (erasure channel) distribution and a Uniform distribution. These distributions are normalized so that they all have the same ratio of the mean to standard deviation. The x-axis is the log of this ratio. This captures the fact that the ratio of the mean to the standard deviation is the only parameter that matters for the second-moment control capacity. As per Corollary 5.6.7 the second moment capacities for all three line up exactly. We see that the Shannon sense capacity for both the Gaussian and the uniform are larger than the second-moment capacity as expected. The Shannon capacity for the Bernoulli distribution is infinity since it contains an atom. The zero-error capacity for the Gaussian channel is zero because it is unbounded. The zero-error Bernoulli channel has zero-error capacity zero because it has an atom at 0. The Uniform distribution follows the zero-error capacity line for bounded distributions, and has slope 1.

Notice that as the ratio of mean to standard deviation goes to infinity, all of the lines approach slope 1. We conjecture that in this "high SNR" regime, this ratio is essentially

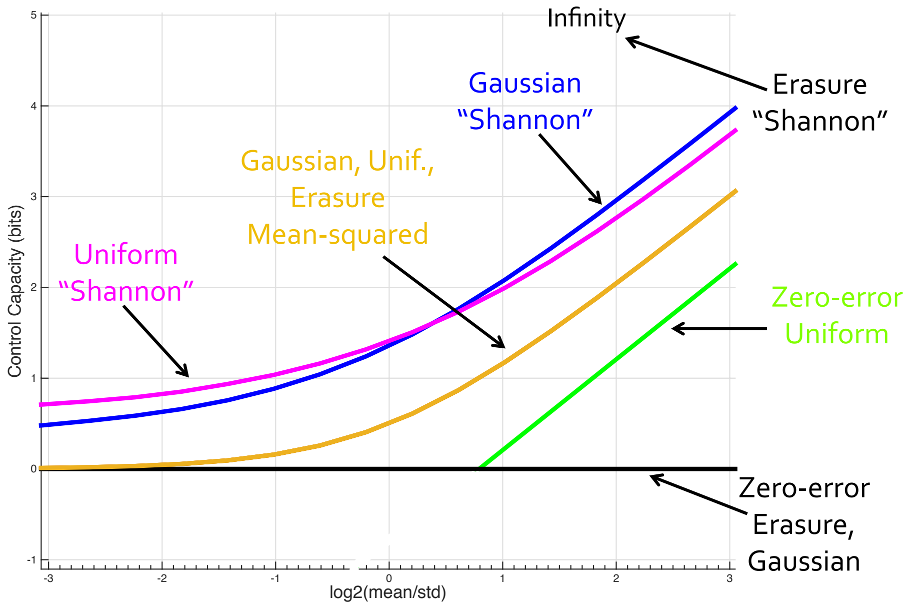


Figure 5.3: Examples of control capacity

what dictates the scaling of control capacity. This is predicted by the carry-free models since the capacity in both the zero-error and Shannon senses depends only on the number of deterministic bits in the control channel gain $g_{det} - g_{ran}$.

Fig. 5.4 allows us to explore the behavior of η-th moment capacities for the same three channels. The plot presents the η-th moment capacities for Gaussian, Erasure and Uniform control channels. We chose the three distributions such that their second-moment capacity is 2, and all three curves intersect there. As $\eta \to 0$, the curves approach the Shannon capacities and as $\eta \to \infty$ the curves asymptote at the zero-error control capacity. The results in this Chapter help us characterize this entire space, while previously only the $(2, 2)$ point was really known.

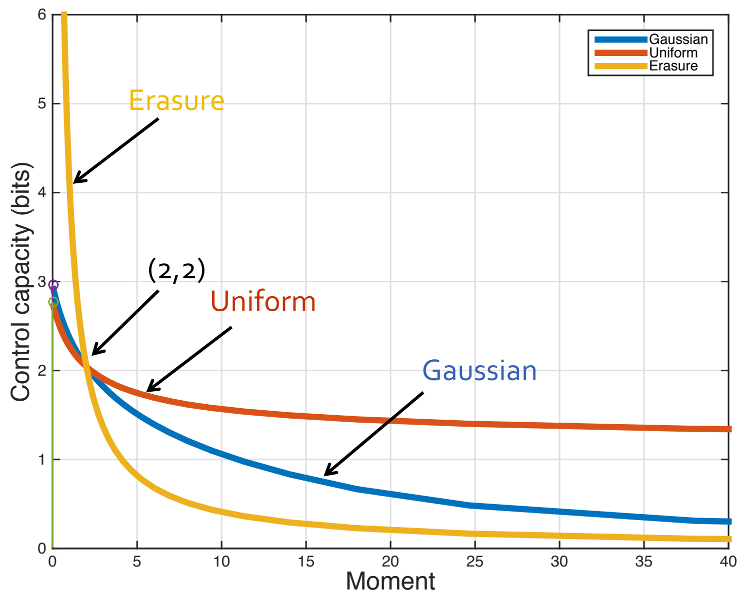


Figure 5.4: The relationship between the different moment-senses of control capacity. For the Uniform, the zero-error control capacity is 1.2075 which is the asymptote as $\eta \to \infty$. As $\eta \to 0$ the η-th moment capacity converges to the Shannon sense. Here the Shannon control capacity for the Uniform is 2.7635 and for the Gaussian is 2.9586, which are the two small points seen on the extreme left (i.e. green and purple on the y-axis).

5.8 Appendix

Lemma 5.4.6.

$$\min_{k\in\mathbb{R}} \max_{B\in[b_1,b_2]} \left|1 + B\cdot k\right| = \begin{cases} \left|\frac{b_2-b_1}{b_2+b_1}\right| & \text{if } \ 0 \notin [b_1,b_2], \ i.e. \ 0 < b_1 < b_2 \ or \ b_1 < b_2 < 0. \\ 1 \ \ if \ \ 0 \in [b_1,b_2] \ i.e. \ b_1 \le 0 \ and \ b_2 > 0, \ \ or \ b_1 < 0 \ and \ b_2 \ge 0. \end{cases} \tag{5.66}$$

This implies that if B is a random variable with essential support $[b_1, b_2]$, essential infimum b_1, essential supremum b_2, then $\forall \epsilon > 0$,

$$P\left(\log\left|1 + B\cdot k\right| < \log\left|\frac{b_2 - b_1}{b_1 + b_2}\right| - \epsilon\right) < 1 \ if \ 0 \notin [b_1, b_2]. \tag{5.67}$$

and

$$P\left(\log\left|1 + B\cdot k\right| < -\epsilon\right) < 1 \ if \ 0 \in [b_1, b_2]. \tag{5.68}$$

Proof:

- Case 1) $0 < b_1 < b_2$ First, we show that the bound is achievable. Choose $k = -\left(\frac{b_1+b_2}{2}\right)^{-1}$. Then, $\left|1 + B \cdot k\right|$ is maximized at $B = b_2$.

$$\left|1 + b_2 \cdot \left(\frac{b_1 + b_2}{2}\right)^{-1}\right| = \frac{b_2 - b_1}{b_1 + b_2} \tag{5.193}$$

Now, suppose $\exists k \in \mathbb{R}$ such that $\forall B \in [b_1, b_2]$,

$$|1 + B \cdot k| < \left|\frac{b_2 - b_1}{b_1 + b_2}\right| = \frac{b_2 - b_1}{b_1 + b_2}. \tag{5.194}$$

This implies:

$$|1 + b_1 \cdot k| < \frac{b_2 - b_1}{b_1 + b_2} \tag{5.195}$$

$$\Rightarrow \frac{b_1 - b_2}{b_1 + b_2} < 1 + b_1 \cdot k < \frac{b_2 - b_1}{b_1 + b_2} \tag{5.196}$$

$$\Rightarrow \frac{-2b_2}{b_1 + b_2} < b_1 \cdot k < \frac{-2b_1}{b_1 + b_2} \tag{5.197}$$

$$\Rightarrow k < \frac{-2}{b_1 + b_2}. \tag{5.198}$$

Also we have,

$$|1 + b_2 \cdot k| < \frac{b_2 - b_1}{b_1 + b_2} \tag{5.199}$$

$$\Rightarrow \frac{b_1 - b_2}{b_1 + b_2} < 1 + b_2 \cdot k < \frac{b_2 - b_1}{b_1 + b_2} \tag{5.200}$$

$$\Rightarrow \frac{-2b_2}{b_1 + b_2} < b_2 \cdot k < \frac{-2b_1}{b_1 + b_2} \tag{5.201}$$

$$\Rightarrow k > \frac{-2}{b_1 + b_2}, \tag{5.202}$$

which gives us a contradiction. Hence, $\forall k \in \mathbb{R}$, $\exists B \in [b_1, b_2]$,

$$|1 + B \cdot k| \geq \left|\frac{b_2 - b_1}{b_1 + b_2}\right|. \tag{5.203}$$

Given the minmax of $|1 + B \cdot k|$ is $\left|\frac{b_2 - b_1}{b_1 + b_2}\right|$, then for any bounded random variable B with essential support on $[b_1, b_2]$ by the definition of the essential supremum b_2 and the essential infimum b_1 we have:

$$P\left(\log\left|1 + B \cdot k\right| < \log\left|\frac{b_2 - b_1}{b_1 + b_2}\right| - \epsilon\right) < 1. \tag{5.204}$$

- Case 2) $b_1 < b_2 < 0$ This follows similarly to Case 1.
- Case 3) $b_1 \leq 0 < b_2$

  To show the bound is achievable, choose $k = 0$. Then $|1 + Bk| = 1$ for all B.

  Suppose there exists $k \in \mathbb{R}$ such that $\forall B \in [b_1, b_2]$,

  $$|1 + B \cdot k| < 1 \tag{5.205}$$

  This implies:

  $$|1 + b_1 \cdot k| < 1 \tag{5.206}$$
  $$\Rightarrow -1 < 1 + b_1 \cdot k < 1 \tag{5.207}$$
  $$\Rightarrow -2 < b_1 \cdot k < 0 \tag{5.208}$$

  - Case 3a) $b_1 = 0$. Then we have a contradiction.
  - Case 3b) $b_1 < 0$

    Since $b_1 \cdot k < 0$, we must have $k > 0$. Now, we also know that

  $$|1 + b_2 \cdot k| < 1 \tag{5.209}$$
  $$\Rightarrow -1 < 1 + b_2 \cdot k < 1 \tag{5.210}$$
  $$\Rightarrow -2 < b_2 \cdot k < 0 \tag{5.211}$$

  $b_2 \cdot k < 0$ implies $k < 0$, which again gives us a contradiction.

  Hence, $\forall k \in \mathbb{R}$, $\exists B \in [b_1, b_2]$,

  $$|1 + B \cdot k| \geq 1. \tag{5.212}$$

  Similarly, for any bounded random variable B with essential support on $[b_1, b_2]$ by the definition of the essential supremum b_2 and the essential infimum b_1 we have:

  $$P\left(\log\left|1 + B[n] \cdot k\right| < -\epsilon\right) < 1. \tag{5.213}$$

- Case 4) $b_1 < 0 \leq b_2$ This is similar to case 3.

■

Chapter 6

The value of side information in control

6.1 Introduction

The advent of networked control systems and stochastic uncertainty models has led to a real need to have a theory capable of dealing with side information in control. As just one example, control theorists are interested in knowing how networked control systems behave with or without acknowledgements of dropped packets since this is relevant for choosing among practical protocols like TCP vs. UDP [101]. Acknowledgements are a kind of side information about control channel state, but as of now, there is no theoretical guidance for how to think about them in a principled way.

Fortunately, side information has long been studied in information theory. There are two sources of uncertainty in point-to-point communication. The first is uncertainty about the system itself, i.e. uncertainty about the channel state, and the second is uncertainty about the message bits in the system. Information theory has looked at both channel-state side information in the channel coding setting and source side information in the source coding setting. Wolfowitz was among the first to investigate channel capacity when the channel state was time-varying [116]. Goldsmith and Varaiya [46] build on this to characterize how channel-state side information at the transmitter can improve channel capacity. Waterfilling (of power) in time is optimal. The time at which side information is received clearly also matters. For instance, in the absence of a probability model for the channel, waterfilling is only possible if all future channel realizations are known in advance[1]. Caire and Shamai [18] provided a unified perspective to understand both causal and imperfect CSI. Lapidoth and Shamai quantify the degradation in performance due to channel-state estimation errors by the receiver [64]. Medard in [72] examines the effect of imperfect channel knowledge on capacity for channels that are decorrelating in time and Goldsmith and Medard [45] further

[1]With a probability model on the channel, the encoder can simulate the future channel realizations, and hence it has an idea of whether the current realization is good or bad.

analyze causal side information for block memoryless channels. Their results recover Caire and Shamai [18] as a special case.

The impact of side information has also been studied in multi-terminal settings. For example, Kotagiri and Laneman [62] consider a helper with non-causal knowledge of the state in a multiple access channel. Finally, there is the surprising result by Maddah-Ali and Tse [67]. They showed that in multi-terminal settings stale channel state information at the encoder can be useful even for a memoryless channel. It can enable a retroactive alignment of the signals. Such stale information is completely useless in a point-to-point setting.

Then there are the classic Wyner-Ziv and Slepian-Wolf results for sourcing coding with source side information that are found in standard textbooks [25]. Pradhan et al. [90] showed that even with side information, the duality between source and channel coding continues to hold: this is particularly interesting given the well-known parallel between source coding and portfolio theory. Source coding with fixed-delay side information can be thought of as the dual problem to channel coding with feedback [68].

Another related body of work looks at the uncertainty in the distortion function as opposed to uncertainty of the source[2]. Rate-distortion theory serves as a primary connection between information theory and control. The distortion function quantifies the importance of the various bits in the message, and in a sense confers a meaning upon them. Uncertainty regarding the distortion function is a way to model uncertainty of meaning. This relates to the toy example we discuss in Section 6.6. Martinian et al. [69] quantified the value of side information regarding the distortion function used to evaluate the decoder [70, 69].

Moving beyond communication, the MMSE dimension looks at the value of side information in an estimation setting. In a system with only additive noise, Wu and Verdu show that a finite number of bits of side information regarding the additive noise cannot generically change the high-SNR scaling behavior of the MMSE [120]. Portfolio theory also gives us an understanding of side information. In the context of horse races, if each race outcome is distributed according a random variable X, then the mutual information between X and Y, $I(X;Y)$, measures the gain in the doubling rate that the side information Y provides the gambler [60]. Directed mutual information captures exactly the causal information that is shared between two random processes. This connection has been made explicit for portfolio theory in [88, 87] by showing that the directed mutual information $I(X^n \to Y^n)$ is the gain in the doubling rate for a gambler due to causal side information Y^n. Of course, directed mutual information is central to control and information theory as the appropriate measure of the capacity of a channel with feedback [109].

A series of works have examined the value of information in stochastic control problems with *additive* disturbances. For a standard LQG problem, Davis [27] defines the value of information as the control cost reduction due to using a clairvoyant controller that has knowledge of the complete past and future of the noise. Rockafellar and Wets [95] first considered this issue for discrete-time finite-horizon stochastic optimization problems, and

[2]A dual notion on the channel coding side is the channel cost function, i.e. the power that it takes per input symbol [38]. What if the encoder does not know this perfectly?

then Dempster [28] and Flam [35] extended the result for infinite horizon problems. Finally, Back and Pliska [9] defined the shadow price of information for continuous time decision problems as well.

This chapter extends control capacity as developed in Chapter 5 to models that include side information. We explore the value of causal side information for control systems in models that involve parameter uncertainty, where these parameters have an i.i.d. character to them. This chapter also looks at the notion of the "tolerable growth" for a control system. We can think of the tolerable growth of a system as a function of both the system and the control strategy that is applied. The next chapter extends these ideas to non-causal side information.

6.2 A carry-free example

As always, we start with a motivating carry-free example for the value of side information.

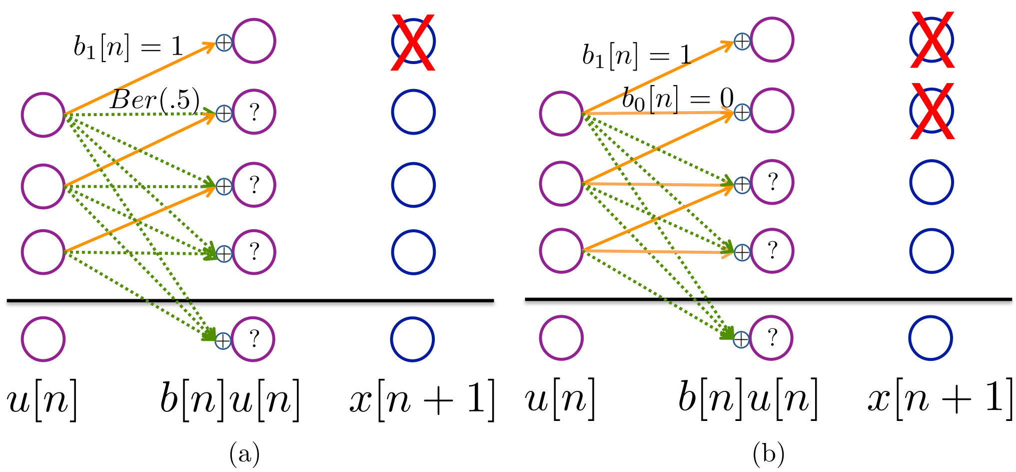


Figure 6.1: This system has the highest deterministic link at level $g_{det} = 1$ and the highest unknown link at $g_{ran} = 0$. Bits $b_{-1}[n], b_{-2}[n], \cdots$ are all random Bernoulli-($\frac{1}{2}$). As a result the controller can only influence the top bits of the control going in, and can only cancel one bit of the state. If one extra bit b_0 were known to the controller, it could cancel a second bit of the state.

In Figure 6.1, we consider a simple bit-level carry-free model that is the counterpart of system (5.1). Say the control gain $B[n]$ has one deterministic bit, so that $g_{det} = 1$, but all lower bits are random Bernoulli-$\frac{1}{2}$ bits. Then the controller can only reliably cancel $1 - 0 = 1$ bits of the state each time. The difference between the level of the deterministic bits and the level of the random bits is what determines the number of reliably controllable bits. If the value of the bit at level 0, i.e. b_0, were also known, then we could tolerate a growth through

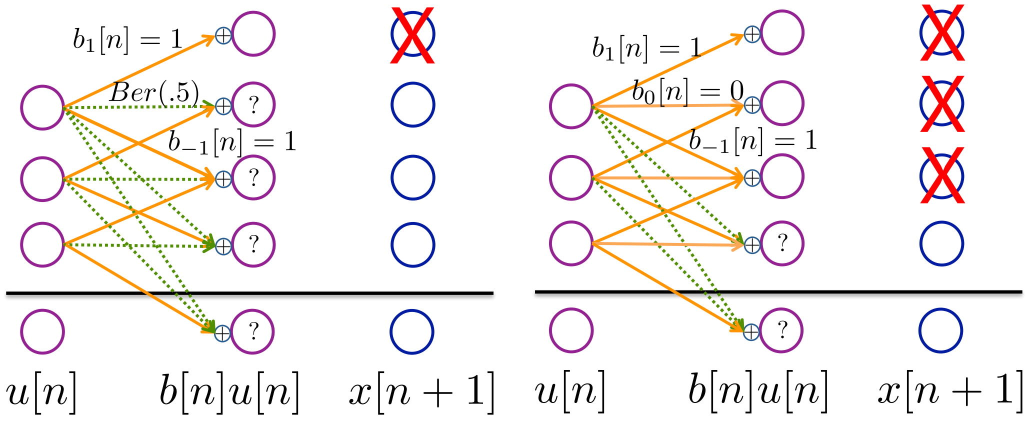


Figure 6.2: Consider the following gain for the controller in (a): $b_1[n] = 1, b_{-1}[n] = 1$ are deterministically known, but all other links are Bernoulli-($\frac{1}{2}$). Only a gain of $\log a = 1$ can be tolerated in this case. Now, say side information regarding the value of $b_0[n]$ is received as in (b). This suddenly buys the controller not just one, but two bits of growth.

a of two bits at a time. We can think of this as the value of the side information b_0 for this problem.

6.3 A carry-free counterexample

In the portfolio theory literature, it is known that the maximum increase in doubling rate due to side information Z for a set of stocks distributed as T is upper bounded by $I(T; Z)$. With our observation about deterministic models it is tempting to conjecture that "a bit buys a bit" and a similar bound holds for the value of information in control systems. However, we see that the following counterexample rejects this conjecture. Consider the carry-free model in Fig. 6.2. Here $u[n]$ is the control action, and $b[n]u[n] = z[n]$ is the control effect. In Fig. 6.2(a) the uncertainty in $b_0[n]$ does not allow the controller to utilize the knowledge that $b_{-1}[n] = 1$ and arbitrarily set the bits of $b[n]u[n]$. However, one bit of information $b_0[n]$ in Fig. 6.2(b), lets the controller buy two bits of gain in the tolerable growth rate as explained in the caption.

This carry-free model represents a real system where $B[n]$ is drawn from a mixture of disjoint uniform distributions, as in Fig. 6.3. The first most significant bit and the third most significant bit are known, but the second most significant bit is not known. The first bit tells us whether $B[n]$ comes from group A and B or group C and D. The third bit only discriminates to the level that is shown in Fig. 6.4, and so the exact mode of the distribution is still unknown to the controller. The second bit finally lowers the variance, as in Fig. 6.5.

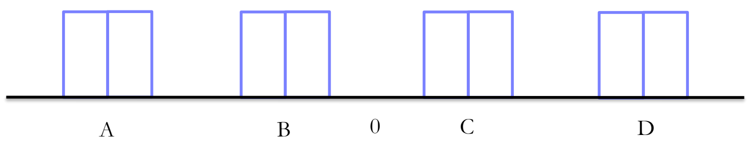


Figure 6.3: The continuous distribution that represents the uncertainty in the $B[n]$ for the carry-free model.

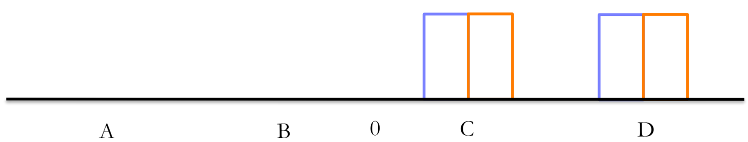


Figure 6.4: The first and the third bit together tell the controller that $B[n]$ comes from one of the orange parts of the distribution. Since there are two orange sections that are far away, the effective variance of $B[n]$ is not reduced.

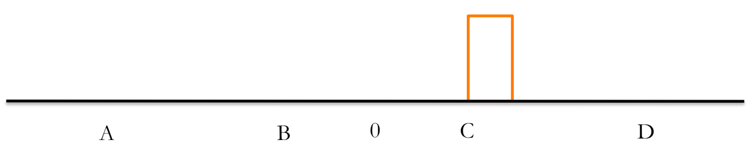


Figure 6.5: Once the second bit is also known, the fact that the controller already knew the third bit becomes useful. Now $B[n]$ comes from a very tight section of the distribution with low variance, which helps to control the system.

A communication aside: If the problem was that of pure communication, we could still decode two bits of information about $u[n]$ from $b[n]u[n] = z[n]$. See Fig. 6.6. Let u_2 and u_0 be the information carrying bits, and set $u_1 = 0$ to zero. Then $z_3 = u_2$ and $z_1 = u_2 + u_0$. With two equations and two unknowns, both $u_2[n]$ and $u_0[n]$ can be recovered at the decoder. In the control problem, this is not possible because of the contamination that is introduced at $z_2 =?$. While communication systems can choose which bits contain relevant information, control systems do not have that flexibility. A bit at a predetrmined position must be cancelled or moved by the control action.

In the case of portfolio theory, it is possible to hedge across uncertainty in the system and get "partial-credit" for uncertain quantities. This is not possible in communication and control systems since it is not possible to hedge a control signal in the same way one can hedge a bet.

Another communication aside: It seems that the "commitment" challenge that is faced by control can also be seen in communication systems, where it is also not possible to hedge

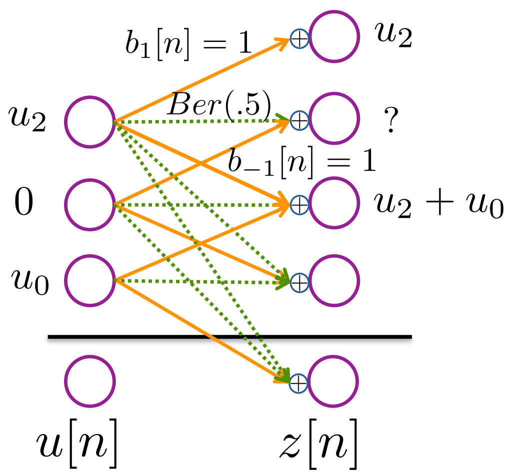


Figure 6.6: With two known bits on the gain, the decoder can decode two additional bits about the message $u[n]$ from the received signal $z[n]$. However, these bits are decoded at specific positions. It is not possible to get information out of the received signal position z_2.

across realizations. Consider a "compound" channel made of two R-bit channels A and B but with distinct inputs, so only one can be used at a time. The message sent across one of the channels is randomly erased with probability 0.5. In this case, one bit of side information about which channel is to be erased can buy us more than a bit: we get $\frac{R}{2}$ bits of message on average.

6.4 Control capacity with side information

Now, we define control capacity with side information. We focus on the Shannon case. The definitions and theorems naturally follow from the definitions of control capacity.

Consider the following system $\mathcal{S}(p_{B[0]}, p_{B[1]}, \cdots, p_{T[0]}, p_{T[1]}, \cdots)$,

$$\begin{aligned} X[n+1] &= X[n] + B[n]U[n] \\ Y[n] &= X[n] \end{aligned} \tag{6.1}$$

$T[i]$ is the side information received by the controller at time i. Now, the control signal $U[n]$ can causally depend on $Y[i], 0 \leq i \leq n$ as well as on the side information $T[i], 0 \leq i \leq n$. Let $\sigma(Y_0^n, T_0^n)$ be the sigma-algebra generated by the observations and the side information. Then $U[n]$ can be a function of these random variables, i.e. $U[n] \in \sigma(Y_0^n, T_{0^n})$. The pairs $(B[i], T[i]), 0 \leq i \leq n$ are generated i.i.d. $p_{B,T}(\cdot, \cdot)$, such that the marginal of $p_{B|T}(\cdot)$ does not have any atoms. $X[0]$ is a random variable with density $p_{X[0]}(\cdot)$ and no atom at 0.

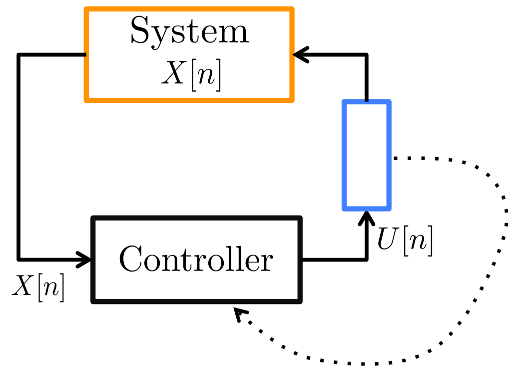


Figure 6.7: The controller has access to side information about the actuation channel.

Definition 6.4.1. The Shannon control capacity with causal side information of the system $\mathcal{S}(p_{B[0]}, p_{B[1]}, \cdots, p_{T[0]}, p_{T[1]}, \cdots)$ is given by

$$C_{sh}(\mathcal{S}|T) = \lim_{n\to\infty} -\min_{U_0^{n-1}(\cdot) \text{ s.t. } U[i](\cdot)\in\sigma(Y_0^i, T_0^i)} \frac{1}{n}\mathbb{E}\left[\log\left|\frac{X[n]}{X[0]}\right|\right]. \tag{6.2}$$

Theorem 6.4.2. *Consider the system $\mathcal{S}(p_B, p_T)$, parameterized the distribution $p_B(\cdot)$ for the i.i.d. $B[n]$s, where $p_B(\cdot)$ is continuous distribution with no atoms, and by the distribution of the side information $p_{T[i]}(\cdot)$ of the i.i.d $T[n]$s. Let $p_{B|T}(\cdot)$ be such that it also contains no atoms, i.e. $T[n]$ can only give a finite number of bits of side information. Then, the Shannon control capacity of the system is given by:*

$$C_{sh}(\mathcal{S}|T) = \mathbb{E}_T\left[\max_k \mathbb{E}\left[-\left|1 + B\cdot k\right|\middle|T\right]\right] \tag{6.3}$$

where $(B, T) \sim p_{B,T}(\cdot,\cdot)$.

Proof: The proof follows from the same argument as in Thm. 5.5.4). Note that Lemma 5.5.5 still applies in this case using the conditional distribution $p_{B|T}(\cdot)$.
Achievability:

The first inequality below follows from the achievability argument in Thm. 5.5.4 following eq. (5.138) to eq. (5.144).

$$\lim_{n\to\infty} \quad -\min_{U_0^{n-1}(\cdot)\text{s.t. } U[i](\cdot)\in\sigma(Y_0^i,T_0^i)} \frac{1}{n}\mathbb{E}\left[\mathbb{E}\left[\log\left|\frac{X[n]}{X[0]}\right|\right]\right] \tag{6.4}$$

$$\geq \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} -\min_{U[i](\cdot)\in\sigma(Y_0^i,T_0^i)} \max_{X[i]\neq 0} \mathbb{E}\left[\log\left|\frac{X[i+1]}{X[i]}\right|\right] \tag{6.5}$$

$$= \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} -\min_{U[i](\cdot)\in\sigma(Y_0^i,T_0^i)} \max_{X[i]\neq 0} \mathbb{E}_{B[i],T[i]}\left[\log\left|\frac{X[i+1]}{X[i]}\right|\right] \tag{6.6}$$

(since the $(B[i],T[i])$ pairs are independent over time we can drop the history in the expectation)

$$= \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} -\min_{U[i](\cdot)\in\sigma(Y_0^i,T_0^i)} \max_{X[i]\neq 0} \mathbb{E}_{T[i]}\left[\mathbb{E}\left[\log\left|\frac{X[i+1]}{X[i]}\right|\middle|T[i]\right]\right] \tag{6.7}$$

$$= \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}_{T[i]}\left[\max_{k(T[i])} \mathbb{E}\left[-\log\left|1+B[i]\cdot k\right|\middle|T[i]\right]\right] \tag{6.8}$$

(Lemma 5.5.5, now applied with conditioning, k can be a function of $T[i]$)

$$= \mathbb{E}_{T[i]}\left[\max_{k} \mathbb{E}\left[-\log\left|1+B[0]\cdot k\right|\middle|T[i]\right]\right] \tag{6.9}$$

Converse: The first inequality below follows from the converse argument in Thm. 5.5.4 following eq. (5.147) to eq. (5.153).

$$\lim_{n\to\infty} \quad - \min_{U_0^{n-1}(\cdot)\text{s.t. } U[i](\cdot)\in\sigma(Y_0^i,T_0^i)} \frac{1}{n}\mathbb{E}\left[\mathbb{E}\left[\log\left|\frac{X[n]}{X[0]}\right|\right]\right] \tag{6.10}$$

$$\leq \lim_{n\to\infty} \quad - \min_{U[n](\cdot)\in\sigma(Y_0^n,T_0^n)} \frac{1}{n} \min_{X[n-1]\neq 0} \mathbb{E}\left[\log\left|\frac{X[n]}{X[n-1]}\right|\right] + - \min_{U_0^{n-2}\text{s.t. } U[i](\cdot)\in\sigma(Y_0^i,T_0^i)} \frac{1}{n}\mathbb{E}\left[\log\left|\frac{X[n-1]}{X[0]}\right|\right] \tag{6.11}$$

$$\leq \lim_{n\to\infty} \quad - \min_{U[n](\cdot)\in\sigma(Y_0^n,T_0^n)} \frac{1}{n} \min_{X[n-1]\neq 0} \mathbb{E}_{T[n-1]}\left[\mathbb{E}\left[\log\left|\frac{X[n]}{X[n-1]}\right|\middle|T[n-1]\right]\right] + \\ - \min_{U_0^{n-2}\text{s.t. } U[i](\cdot)\in\sigma(Y_0^i,T_0^i)} \frac{1}{n}\mathbb{E}\left[\log\left|\frac{X[n-1]}{X[0]}\right|\right] \tag{6.12}$$

$$= \lim_{n\to\infty} \frac{1}{n}\mathbb{E}_{T[n-1]}\left[\max_k \mathbb{E}\left[-\log\left|1+B[n-1]\cdot k\right|\middle|T[n-1]\right]\right] + \\ - \min_{U_0^{n-2}\text{s.t. } U[i](\cdot)\in\sigma(Y_0^i,T_0^i)} \frac{1}{n}\mathbb{E}\left[\log\left|\frac{X[n-1]}{X[0]}\right|\right] \tag{6.13}$$

(Lemma 5.5.5, now applied with conditioning)

$$= \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}_{T[i]}\left[\max_k \mathbb{E}\left[-\log\left|1+B[i]\cdot k\right|\middle|T[i]\right]\right] \tag{6.14}$$

(by induction, since the $(B[i], T[i])$ pairs are i.i.d.)

$$= \mathbb{E}_{T[0]}\left[\max_k \mathbb{E}\left[-\log\left|1+B[0]\cdot k\right|\middle|T[0]\right]\right] \tag{6.15}$$

■

Remark 6.4.3. For scalar systems like the one considered above, non-causal side information about the actuation channel would actually not provide any capacity gains. Receiving $T[2]$ at timestep-1 over timestep-2 would not change the system performance.

Remark 6.4.4. These basic ideas on side information could be extended to both the zero-error and η-th moment senses of capacity. For the zero-error case we can replace the expectation with a min. For the second-moment case, we must deal with the added variability due to the side information. Notice that the results parallel the Shannon communication channel capacity with side information about the channel state available at *both* the transmitter and the receiver.

As an example, we plot the change in Shannon control capacity with side information for a set of actuation channels with a Uniform uncertainty distribution in Fig. 6.8. It turns out that the critical parameter to calculate control capacity (as suggested by the carry-free models) is the ratio of the mean to the standard deviation of the distribution. All distributions here have mean 1. Consider $B \sim \text{Uniform}[b_1, b_2]$. Then, one bit of side information divides this interval into two halves, and tells the controller whether the realization of B is in $[b_1, \frac{b_1+b_2}{2}]$ or in $(\frac{b_1+b_2}{2}, b_2]$. These intervals are further subdivided with more bits.

The figure looks at four different ratios of the mean to the standard deviation $100, 4, \frac{1}{4}, \frac{1}{100}$. First, let us look at the green curve for the distribution with mean to standard deviation ratio 100. As the number of bits of side information increases the slope of this curve approaches 1, but it is slightly below 1 close to 0. The same is true for the pink line with mean to standard deviation ration 4. The cases with mean to standard deviation less than 1 behave differently. The yellow line, for ratio $\frac{1}{4}$, has a slope slightly greater than 1 close to zero. This distribution is a Uniform on $[-5.92, 7.92]$. Here, half the time, the first bit of information reveals perfectly the sign of the distribution. So the capacity can increase by more than one bit. The same holds when the ratio is $\frac{1}{100}$, with an even sharper slope close to 0. As the number of bits increase, again the slope approaches 1. The carry-free model predicted this behavior — the value of a bit can be more than a bit!

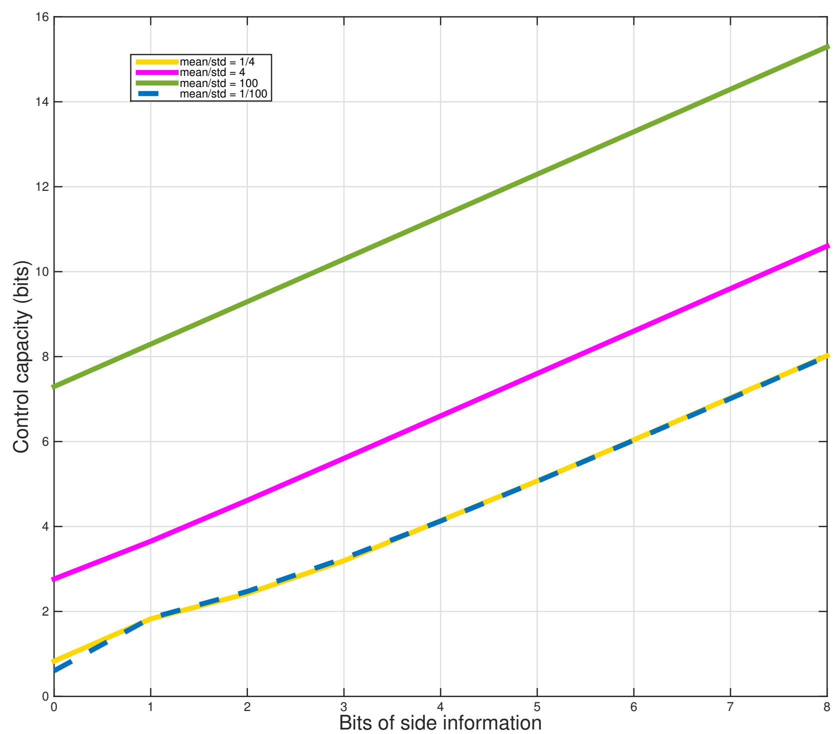


Figure 6.8: This plot shows the increase in Shannon control capacity of channels with a uniform distribution for distributions with mean 1 and mean/standard deviation ratio as in the legend.

6.5 Tolerable system growth

This section explores control capacity through the concept of the tolerable system growth, i.e. the maximal system gain that can be tolerated while maintaining stabilizability. This is directly related to the theorems in Chapter 5, (Thm. 5.4.3, Thm. 5.5.2, and Thm. 5.6.3), that show the operational meaning of control capacity for non-unit system gains for different senses of stability. The developments here could in principle be used to generalize the ideas of control capacity to more general systems than the scalar LTI systems in Chapter 5.

6.5.1 Control

Consider a real-valued control system $\mathcal{S}^a$, with state $X[n]$, control $U[n]$ and observation $Y[n]$ at time n as below

$$\begin{aligned} X[n+1] &= a \cdot f(X[n], U[n], \mathcal{T}[n]), \\ Y[n] &= g(X[n], \mathcal{T}[n]). \end{aligned} \tag{6.16}$$

Let $\mathcal{T}[n]$ be the set of random variables associated with the system at time n. Let $F_{\mathcal{T}[n]}$ be the set of distributions associated with them, and we assume these are known to the controller. f and g are fixed, known, deterministic functions of $X[n]$ and $U[n]$. a is a scalar, known constant. The control strategy is a function $U[n] : \sigma(Y_0^n) \to \mathbb{R}$.

For instance, for the system (5.1), $\mathcal{T}[n] = \{B[n]\}$ and $F_{\mathcal{T}[n]}$ is effectively $\{F_B\}$ at each n since $B[n]$ are i.i.d..

Logarithmic stability

Corresponding to Shannon control capacity, we define a one-step logarithmic decay rate.

Definition 6.5.1. For the system (6.16) with state $X[n]$, control strategy $U[n]$ and system randomness $F_{\mathcal{T}[n]}$ at time n as, the one-step logarithmic decay is defined as:

$$G_{\mathcal{S}^a}(X[n], U[n], F_{\mathcal{T}[n]}) \stackrel{\text{def}}{=} \mathbb{E}\left[-\log \frac{||X[n+1]||}{||X[n]||}\right]. \tag{6.17}$$

The expectation is over the randomness $F_{\mathcal{T}[n]}$.

For example, for the system from Chapter 5, in eq. (5.1), the function $f(X[n], U[n], B[n]) = X[n] + B[n]U[n]$ and $g(X[n]) = X[n]$. The random variables $B[n] \sim \text{Uniform}[b_1, b_2]$, and $U[n] = -\frac{2}{b_1+b_2}X[n]$, $G_{\mathcal{S}^a}(X[n], U[n], F_B) = \log|a| + \mathbb{E}\log|1 - \frac{2B[n]}{b_1+b_2}|$.
Related to these rates we can define the average logarithmic decay rate of a system.

Definition 6.5.2. The average logarithmic decay rate of the system (6.16) is

$$\bar{G}_{\mathcal{S}^a}(X[0], U_0^\infty, F_{\mathcal{T}_0^\infty}) \stackrel{\text{def}}{=} \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{E}\left[G_{\mathcal{S}^a}(X[i], U[i], F_{\mathcal{T}[i]})\right] \tag{6.18}$$

$$= \lim_{n\to\infty} \frac{1}{n} \mathbb{E}\left[\log \frac{||X[n+1]||}{||X[0]||}\right] \quad \text{(telescoping sum).} \tag{6.19}$$

The expectation inside the sum in eq. (6.18) is over the random state $X[i]$. $X[0]$ is also a random variable.

The connections of this decay rate to different notions of stabilizability are discussed further in the Appendix 6.7.1. In particular, these notions can help us extend the ideas of control capacity to systems with both additive and multiplicative noise.

Remark 6.5.3. The maximal average decay rate, $G^*_{\mathcal{S}^a}(F_{\mathcal{T}_0^\infty})$, defined as the decay rate of the system when the optimal control is applied, is the same as the Shannon control capacity of the system.

$$G^*_{\mathcal{S}^a}(F_{\mathcal{T}_0^\infty}) = C_{sh}(\mathcal{S}^a). \tag{6.20}$$

We hope to use this in future work to generalize the notions of control capacity beyond linear systems.

Now, consider eq. (6.16) with $a = 1$. Call this system $\mathcal{S}^1$.

$$\begin{aligned} X[n+1] &= 1 \cdot f(X[n], U[n], \mathcal{T}[n]), \\ Y[n] &= g(X[n], \mathcal{T}[n]). \end{aligned} \tag{6.21}$$

This system with no explicit growth through a, determines the bottleneck decay rate for the functions f and g.

Definition 6.5.4. The tolerable growth for a system $\mathcal{S}^a$ is given by the maximal average decay rate of the corresponding system $\mathcal{S}^1$:

$$a_0^* = 2^{G^*_{\mathcal{S}^1}(F_{\mathcal{T}_0^\infty})}. \tag{6.22}$$

for system (6.21), i.e. $\log a_0^* = G^*_{\mathcal{S}^1}(F_{\mathcal{T}_0^\infty})$.

The a_0^* establishes the "region of convergence" for the system[3].

We can also define the tolerable growth for a particular strategy U_0^∞, again using $\mathcal{S}^1$ as the base system.

$$a_0(U_0^\infty) = \exp\{\mathbb{E}\left[\bar{G}_{\mathcal{S}^1}(X[0], U_0^\infty, F_{\mathcal{T}_0^\infty})\right]\}. \tag{6.23}$$

[3]We are borrowing the notion of a region of convergence from the theory of Laplace transforms. The Laplace transform of a signal $f(t)$ is defined as $F(s) = \int_{-\infty}^{\infty} e^{-st} f(t)dt$. This integral converges for parameters s that are in the region of convergence of the signal and it is meaningful to talk about other properties of the signal in this domain only.

η-th moment stability

This section defines the tolerable growth rates for the η-th moment of systems, and helps us see why the definition of Shannon control capacity is right notion for the zeroth moment control capacity as well.

Definition 6.5.5. For the system (6.16) with state $X[n]$, control strategy $U[n]$ and system randomness $F_{\mathcal{T}[n]}$ at time n as, define the one-step η-th moment decay as:

$$G_{\mathcal{S}^a,\eta}(X[n], U[n], F_{\mathcal{T}[n]}) \stackrel{\text{def}}{=} -\frac{1}{\eta}\log \mathbb{E}\left[\frac{||X[n+1]||^\eta}{||X[n]||^\eta}\right]. \tag{6.24}$$

The expectation is over the randomness $F_{\mathcal{T}[n]}$.

Definition 6.5.6. The average η-th moment decay rate ($\eta > 0$) of the system (6.16) is

$$\bar{G}_{\mathcal{S}^a,\eta}(X[0], U_0^\infty, F_{\mathcal{T}_0^\infty}) \stackrel{\text{def}}{=} \lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}\left[G_{\mathcal{S}^a,\eta}(X[i], U[i], F_{\mathcal{T}[i]})\right]. \tag{6.25}$$

The expectation inside the sum is over the random state $X[i]$.

Again, the optimal control strategy determines the maximal average η-th moment decay rate, $G^*_{\mathcal{S}^a,\eta}(F_{\mathcal{T}_0^\infty})$, with the expectation over $X[0]$. The max is over all causal control strategies $U[n] = h(Y_0^n)$, where $h(\cdot)$ is some (possibly non-linear) function of the observations till time n.

Definition 6.5.7. The maximal average η-th moment decay rate is given by

$$G^*_{\mathcal{S}^a,\eta}(F_{\mathcal{T}_0^\infty}) \stackrel{\text{def}}{=} \max_{U_0^\infty,\ U[n]=h(Y_0^n)} \mathbb{E}\left[\bar{G}_{\mathcal{S}^a}(X[0], U_0^\infty, F_{\mathcal{T}_0^\infty})\right]. \tag{6.26}$$

The expectation is over $X[0]$ which is a random variable.

We can also define a tolerable growth rate a^*_η for the η-th moment.

Definition 6.5.8. For system (6.21), the maximum growth rate for η-th moment stability is

$$a^*_\eta = 2^{G^*_{\mathcal{S}^1,\eta}(F_{\mathcal{T}_0^\infty})}. \tag{6.27}$$

Finally, we define the η-th moment tolerable growth for a particular control strategy similar to the definition for the logarithmic tolerable growth earlier.

$$a_\eta(U_0^\infty) = 2^{\bar{G}_{\mathcal{S}^1}(X[0], U_0^\infty, F_{\mathcal{T}_0^\infty})} \tag{6.28}$$

As $\eta \to 0$, the tolerable growth for the η-th moment converges to the logarithmic tolerable growth. Hence, the logarithmic tolerable growth and the corresponding decay rate essentially correspond to the "zeroth" moment of the system. The idea here is that the logarithmic growth rate is similar to the geometric mean of the system decay rate over time.

Theorem 6.5.9. $\lim_{\eta\to 0} a_\eta(U_0^\infty) = a_0(U_0^\infty)$

The proof is in the Appendix. 6.7.2.

6.5.2 Estimation

We also define a corresponding decay rate for estimation. Consider the system $\mathcal{S}^a$:

$$\begin{aligned} X[n+1] &= a \cdot f(X[n], \mathcal{T}[n]), \\ Y[n] &= g(X[n], \mathcal{T}[n]). \end{aligned} \tag{6.29}$$

Definition 6.5.10. Define the one-step logarithmic error decay rate for an estimator $\widehat{X}[n]$ that is a function of $Y[1], Y[2] \cdots Y[n]$

$$G_{\mathcal{S}^a}(\widehat{X}[n], F_{\mathcal{T}[n]}) \stackrel{\text{def}}{=} \mathbb{E}\left[-\log \frac{||X[n+1] - \widehat{X}[n+1]||}{||X[n] - \widehat{X}[n]||} \right]. \tag{6.30}$$

The average logarithmic decay rate can then be defined as in the control case.

Definition 6.5.11. The average logarithmic decay rate for a system $\mathcal{S}^a$ and estimator $\widehat{X}[n]$ is defined as

$$\bar{G}_{\mathcal{S}^a}(X[0], \widehat{X}_0^\infty, F_{\mathcal{T}_0^\infty}) \stackrel{\text{def}}{=} \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{E}\left[G_{\mathcal{S}^a}(X[i], U[i], F_{\mathcal{T}[i]}) \right] \tag{6.31}$$

The expectation inside the sum in eq. (6.18) is over the random state $X[i]$. $X[0]$ is also a random variable.

Finally, the value of side information $Z[n]$ for estimator $\widehat{X}[n]$ is the change in the average decay rate.

Definition 6.5.12. Let $\mathcal{Z}[n]$ be the side information provided to the controller at time n. Then the average value of side information $\mathcal{Z}[n]$ with the control strategy U_0^∞ is

$$\mathbb{E}\left[\bar{G}_{\mathcal{S}^a}(X[0], \widehat{X}_0^\infty, F_{\mathcal{T}_0^\infty | \mathcal{Z}_0^\infty}) - \bar{G}_{\mathcal{S}^a}(X[0], \widehat{X}_0^\infty, F_{\mathcal{T}_0^\infty}) \right]. \tag{6.32}$$

The expectation is taken over the random side information vector $\mathcal{Z}_0^\infty$.

The optimal estimator determines the maximal average decay rate, $G^*_{\mathcal{S}^a}(F_{\mathcal{T}_0^\infty})$, with the expectation over $X[0]$. The max is over all causal estimators $\widehat{X}[n] = h(Y_0^n)$, where $h(\cdot)$ is some (possibly non-linear) function of the observations till time n.

$$G^*_{\mathcal{S}^a}(F_{\mathcal{T}_0^\infty}) \stackrel{\text{def}}{=} \max_{\widehat{X}_0^\infty,\ \widehat{X}[n]=h(Y_0^n)} \mathbb{E}\left[\bar{G}_{\mathcal{S}^a}(X[0], \widehat{X}_0^\infty, F_{\mathcal{T}_0^\infty}) \right]. \tag{6.33}$$

With this we can define the value of side information for estimation (without regard to an estimator) as the change in the maximal growth rate

Definition 6.5.13. Let $\mathcal{Z}[n]$ be the side information provided to the estimator at time n. Then the average value of side information $\mathcal{Z}[n]$ is

$$\mathbb{E}\left[G^*_{\mathcal{S}^a}(F_{\mathcal{T}_0^\infty|\mathcal{Z}_0^\infty}) - G^*_{\mathcal{S}^a}(F_{\mathcal{T}_0^\infty})\right]. \tag{6.34}$$

From arguments in Chapter 4, we know that value of any finite number of bits of side side information bits provided about an unknown continuous random gain $C[n]$ on the observations is 0, even though we saw that information bits about $B[n]$ are valuable for the control problem.

6.6 A toy example

The examples in this dissertation so far have focused on scalar systems. Here we consider a vector example where we can explicitly calculate how the decay rate and tolerable growth change as a function of side information. This problem is actually motivated by the intermittent Kalman filtering [102, 84] and dropped control [30, 101] problems examined in the Chapter 7. We continue to build on this example in Chapter 7.

Here, we focus on uncertainty in the control channel first, and then on uncertainty in the observation channel. The example shows how the dual problems of estimation and control behave differently in the presence of multiplicative noise, even though they have similar behavior when only additive noise is present (e.g. Tatikonda and Mitter [110]).

6.6.1 A spinning controller

Consider the noiseless 2D control system in (6.35)

$$\begin{aligned} \begin{bmatrix} X_1[n+1] \\ X_2[n+1] \end{bmatrix} &= a \begin{bmatrix} \cos\Phi_n & \sin\Phi_n \\ -\sin\Phi_n & \cos\Phi_n \end{bmatrix} \left(\begin{bmatrix} X_1[n] \\ X_2[n] \end{bmatrix} + U[n] \begin{bmatrix} \cos\Theta_n \\ \sin\Theta_n \end{bmatrix} \right), \\ Y[n] &= \begin{bmatrix} X_1[n] \\ X_2[n] \end{bmatrix}. \end{aligned} \tag{6.35}$$

The controller has perfect access to the system state, denoted as a vector using boldface $\mathbf{X}[n] = \begin{bmatrix} X_1[n] & X_2[n] \end{bmatrix}^T$ but is subject to the following limitation: at each time n, the available actuation direction is determined by a random spin. Θ_n is uniformly distributed from $[0, 2\pi]$. Some information about the realized control direction, $\Theta_n = \theta_n$ is revealed to the controller (we call this side information $Z[n]$, and in the complete information case $Z[n] = \theta_n$) before it chooses $U[n]$. The controller may only act along the direction $\begin{bmatrix} \cos\theta_n & \sin\theta_n \end{bmatrix}^T$. The initial state $\begin{bmatrix} X_1[0] & X_2[0] \end{bmatrix}^T$ is drawn randomly according to a known distribution, and the goal is to drive the state to the origin. Thus, if the time-horizon is N, the goal is to minimize:

$$J_N(\mathbf{X}[N+1]) = \mathbb{E}\left[||\mathbf{X}[N+1]||^2\right] = \mathbb{E}\left[X_1[N]^2 + X_2[N]^2\right] \tag{6.36}$$

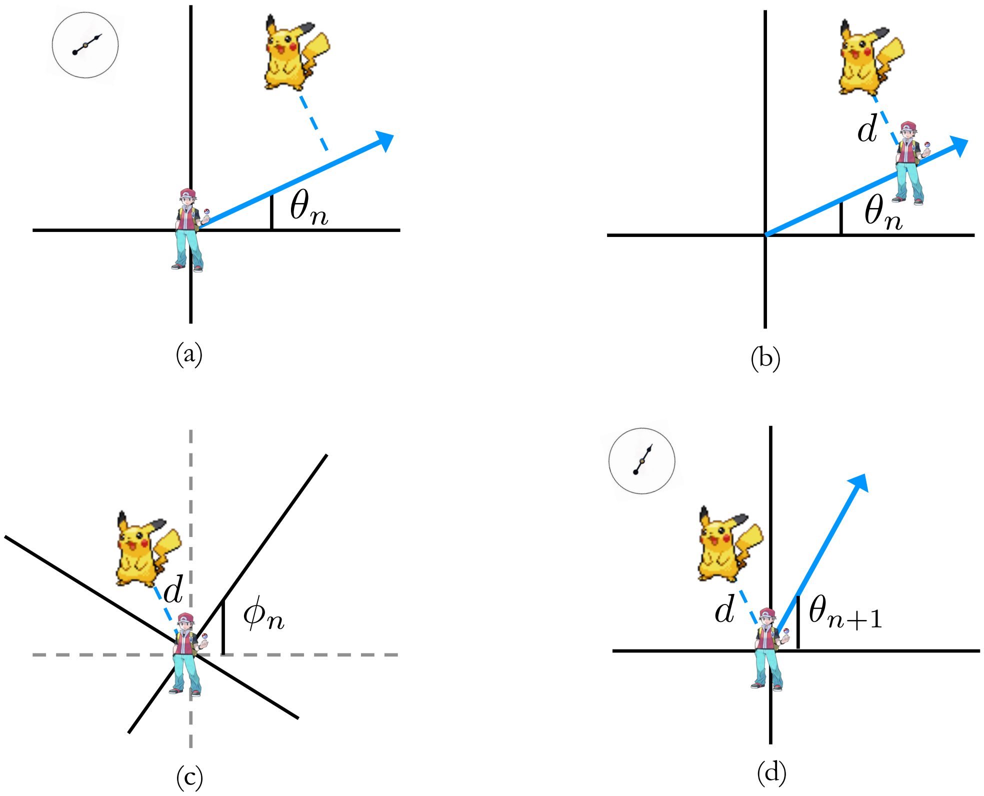


Figure 6.9: The yellow Pikachu figure is the target that the controller must reach. (a) The controller can only move along the direction θ_n. This is randomly chosen by the spinner. (b) The controller choses the control $U[n]$ and moves along the direction θ_n. (c) After this move, the system axes are randomly spun and revealed to the controller. Thus, the distance d between controller and target is maintained, but the angle is not. (d) A new direction θ_{n+1} is chosen by the spinner and the game continues.

After the control acts, the system axes are spun again by the rotation matrix $\begin{bmatrix} \cos \Phi_n & \sin \Phi_n \\ -\sin \Phi_n & \cos \Phi_n \end{bmatrix}$. Φ_n is also randomly and uniformly drawn from $[0, 2\pi]$ at each time and the realization $\Phi_n = \phi_n$ is revealed to the controller at time $n+1$. So $U[n]$ is a function of $Y[n]$, the side information $Z[n] = \theta_n$ and the axis-spin realization ϕ_{n-1}. This is depicted in Fig. 6.9.

The impact of this Φ is that only the distance of the target from the origin is preserved across time. The angle of the state after $U[n]$ is applied is irrelevant since it is subject to the random axis spin. Finally, the scale a is applied.

No information about θ_n

Consider the case where no information about the realization of control direction θ_n is known to the controller. The controller has 0 bits of information about the system randomness. Clearly, in this case the optimal control action is $U[n] = 0$, and no growth a can be tolerated. The system is stable if and only if $a \leq 1$. For the other extreme case, where the controller knows the present control direction perfectly, the following theorem characterizes the optimal strategy.

Complete side information $Z[n] = \theta_n$, Mean-squared decay

Theorem 6.6.1. *The optimal control to minimize the mean-square system state for the system (6.35) is given by the greedy strategy, i.e.* $U^*[n] = -X_1[n]\cos\theta_n - X_2[n]\sin\theta_n$ *for all* n*. The mean-squared tolerable growth,* $a_2^* = \sqrt{2}$*.*

Proof: The proof follows using dynamic programming. Let the time-horizon for the system be $N+1$. Then, the system objective is to minimize the terminal cost, $J_N(\mathbf{X}[N+1])$.

$$J_N(\mathbf{X}[N+1]) = \mathbb{E}\left[X_1[N]^2 + X_2[N]^2\right] \tag{6.37}$$

Let $J_n(\mathbf{X}[n], U[n], \theta_n)$ be the cost incurred at time n in state $X[n]$ when the realized control angle is θ_n and $U[n]$ is the control applied. Then

$$J_n^*(\mathbf{X}[n]) = \mathbb{E}\left[\min_{U[n]} J_n(\mathbf{X}[n], U[n], \theta_n)\right] \tag{6.38}$$

is the optimal cost-to-go at time n. The expectation is over the realization of the random direction θ_n.
Further, since there is no control cost we have:

$$J_n(\mathbf{X}[n], U[n], \theta_n) = J_{n+1}^*(\mathbf{X}[n+1]), \tag{6.39}$$

where the state $\mathbf{X}[n+1]$ is determined by the action in the previous step through

$$||\mathbf{X}[n+1]||^2 = (X_1[n] + U[n]\cos\theta_n)^2 + (X_2[n] + U[n]\sin\theta_n)^2. \tag{6.40}$$

The angle of the state is determined through the random spin Φ_n, but this has no effect on the control cost.
At time N

$$\begin{aligned} J_N^*(\mathbf{X}[N]) &= \mathbb{E}\left[\min_{U[N]} \ J_N(\mathbf{X}[N], U[N], \theta_N)\right] && (6.41)\\ &= \mathbb{E}\left[\min_{U[N]} \ (X_1[N] + U[N]\cos\theta_N)^2 + (X_2[N] + U[N]\sin\theta_N)^2\right] && (6.42) \end{aligned}$$

We know (or can calculate by differentiating with respect to $U[N]$) that the minimizing control is $U^*[N] = -X_1[N]\cos\theta_N - X_2[N]\sin\theta_N$, i.e. the projection of the initial state onto the control direction. The terminal cost-to-go is the magnitude of the error vector after the projection:

$$J_N^*(\mathbf{X}[N]) = \mathbb{E}\left[(X_1[N]\cos\theta_N - X_2[N]\sin\theta_N)^2\right] \tag{6.43}$$

$$= \frac{1}{2}(X_1^2[N] + X_2^2[N]^2). \tag{6.44}$$

Hence $\mathbb{E}\left[\frac{X_1[N+1]^2+X_2[N+1]^2}{X_1[N]^2+X_2[N]^2}\right] = \frac{1}{2}$. Since the optimal cost-to-go is a function only of the state magnitude, the greedy strategy is optimal for all time horizons.

The expected decay at each time step is constant and independent of N so the maximum tolerable growth is $a_2^* = \sqrt{2}$.

■

Complete side information $Z[n] = \theta_n$, logarithmic decay

Theorem 6.6.2. *Consider system (6.35) with perfect information about θ_n at each time n, and apply the optimal control $U[n] = -X_1[n]\cos\theta_n - X_2[n]\sin\theta_n$ (which minimizes the mean-squared state). Then, the logarithmic decay rate is, $G_S(X[n], U[n], \mathcal{F}_\Theta, \mathcal{F}_\Phi) = \log 2$, and the tolerable growth rate is thus $a_0^* = 2$.*

Proof: We know from the proof of Thm. 6.6.1 that when $U[N]$ is applied at time step N, $\mathbb{E}\left[||X[N+1]||^2\right] = \mathbb{E}\left[(X_1[N]\cos\theta_N - X_2[N]\sin\theta_N)^2\right] = \frac{1}{2}(X_1^2[N] + X_2^2[N]^2)$.

So the logarithmic decay rate of the system for is given by

$$\mathbb{E}\left[-\log\sqrt{\frac{|X_1[N+1]|^2 + |X_2[N+1]|^2}{|X_1[N]|^2 + |X_2[N]|^2}}\right] \tag{6.45}$$

$$= \frac{1}{2\pi}\int_0^{2\pi} \frac{1}{2}\log\frac{|X_1[N]\cos\theta - X_2[N]\sin\theta|^2}{|X_1[N]|^2 + |X_2[N]|^2} d\theta \tag{6.46}$$

We can use a computer (Mathematica) to show that this integral evaluates to 1. Hence, $a_0^* = 2$. ■

6.6.2 Partial side information

The symmetric randomness in this example makes it easy to evaluate how the multiplicative uncertainty in the problem affects our ability to control the system. How does the decay rate of the system change with less than perfect information about the control direction θ_n? At what rate can we control the system if $Z[n]$ only has a few bits of information about θ_n and how valuable is this information?

To understand this, we develop a series of finer and finer quantizations for the realization of the system randomness $\Theta_n = \theta_n$, and use these to provide side information $Z[n]$ to the controller at time n. For example, one quantization is $H_1 = [0, \pi]$ and $H_2 = [\pi, 2\pi]$. Say a genie observes the realization θ_n, and reveals only the half-space of the realization, $Z[n] = H_1$ or $Z[n] = H_2$ to the controller. This is effectively 1 bit of partial side information. Two bits of side information $Z[n]$ would correspond to dividing the space into four quadrants, $Q_1 = [0, \frac{\pi}{2}], Q_2 = [\frac{\pi}{2}, \pi], Q_2 = [\pi, \frac{3\pi}{2}]$ and $Q_4 = [\frac{3\pi}{2}, 2\pi]$ and revealing that to the controller. The controller must plan its control action at time n, $U[n]$, based only on these few bits of information. The case considered in the previous section had perfect knowledge of θ_n at time n, and so essentially had infinite bits of side information with $Z[n] = \theta_n$.

The system is the same as that in (6.35)

$$\begin{aligned} \begin{bmatrix} X_1[n+1] \\ X_2[n+1] \end{bmatrix} &= a \begin{bmatrix} \cos\Phi_n & \sin\Phi_n \\ -\sin\Phi_n & \cos\Phi_n \end{bmatrix} \left(\begin{bmatrix} X_1[n] \\ X_2[n] \end{bmatrix} + U[n] \begin{bmatrix} \cos\Theta_n \\ \sin\Theta_n \end{bmatrix} \right), \\ Y[n] &= \begin{bmatrix} X_1[n] \\ X_2[n] \end{bmatrix}. \end{aligned} \tag{6.47}$$

except now the control action $U[n]$ is a function of $Y[n]$ and $Z[n]$, as opposed to $Y[n]$ and θ_n in the earlier setup.

Theorem 6.6.3. *The mean-squared and logarithmic decay rates for for system* (6.35) *with* $k = 0, 1, 2, 3$ *bits of information about the control direction* θ_n *are as in Table. 6.1.*

Table 6.1: System growth as a function of current state-information

		2-nd moment tolerable growth, a_2^*	Logarithmic tolerable growth, a_0^*, lower bound
0 bits		1.00	1.00
1 bit		1.12	1.22
2 bit		1.29	1.61
3 bits		1.37	1.86
∞ bits		$\sqrt{2} = 1.41$	2.00

Proof: We use a dynamic program to calculate the values above. We demonstrate the idea for $k = 2$ bits of side information in Lemma 6.6.4 and Lemma 6.6.5 as representative examples. Calculations for all other cases follow similarly. ■

Lemma 6.6.4. *The mean-squared tolerable growth rate with two bits of side information,* $Z[n] \in \{Q_1, Q_2, Q_3, Q_4\}$, *for the system* (6.47) *is* $a_2^* = 1.29$.

Proof: Our aim is to minimize the terminal cost

$$J_N(\mathbf{X}[N+1]) = \mathbb{E}\left[X_1[N]^2 + X_2[N]^2\right]. \tag{6.48}$$

Then, for every n,

$$J_n^*(\mathbf{X}[n], Z[n]) = \frac{1}{4}\left[\min_{U[n]} \mathbb{E}[J_n(\mathbf{X}[n], U[n], Z[n] = Q_1)] + \min_{U[n]} \mathbb{E}[J_n(\mathbf{X}[n], U[n], Z[n] = Q_2)] + \right.$$
$$\left. \min_{U[n]} \mathbb{E}[J_n(\mathbf{X}[n], U[n], Z[n] = Q_3)] + \min_{U[n]} \mathbb{E}[J_n(\mathbf{X}[n], U[n], Z[n] = Q_4)]\right]. \tag{6.49}$$

We can calculate the optimal control for each quadrant, as the one that minimizes the cost given that θ belongs to the given quadrant.

$$\operatorname*{argmin}_{U_1[n]} \mathbb{E}[J_n(\mathbf{X}[n], U_1[n], Z[n] = Q_1)] \tag{6.50}$$

$$= \frac{2}{\pi}\int_0^{\pi/2} (X_1[n] + U[_1n]\cos\theta)^2 + (X_2[n] + U_1[n]\sin\theta)^2 d\theta \tag{6.51}$$

$$= 2\frac{X_1[n] + X_2[n]}{\pi} \tag{6.52}$$

Similarly, we can compute:

$$\operatorname*{argmin}_{U_2[n]} \mathbb{E}[J_n(X[n], U_2[n], Z[n] = Q_2)] = -2\frac{X_1[n] - X_2[n]}{\pi} \tag{6.53}$$

$$\operatorname*{argmin}_{U_3[n]} \mathbb{E}[J_n(X[n], U_3[n], Z[n] = Q_3)] = -2\frac{X_1[n] + X_2[n]}{\pi} \tag{6.54}$$

$$\operatorname*{argmin}_{U_4[n]} \mathbb{E}[J_n(X[n], U_4[n], Z[n] = Q_4)] = 2\frac{X_1[n] - X_2[n]}{\pi} \tag{6.55}$$

Unlike in the case with perfect side information $Z[n]$, here the optimal control is not allowed to depend on θ_n. Effectively the optimal control is chosen to minimize the cost averaged over all possible directions in the given quadrant.

$$
\begin{aligned}
&J_n^*(\mathbf{X}[n], Z[n]) \\
&= \frac{1}{4}\frac{2}{\pi}\int_0^{\pi/2}\left[\left(X_1[n] - 2\frac{X_1[n]+X_2[n]}{\pi}\cos\theta\right)^2 + \left(X_2[n] - 2\frac{X_1[n]+X_2[n]}{\pi}\sin\theta\right)^2\right] d\theta \\
&+ \frac{1}{4}\frac{2}{\pi}\int_{\pi/2}^{\pi}\left[\left(X_1[n] + 2\frac{X_1[n]-X_2[n]}{\pi}\cos\theta\right)^2 + \left(X_2[n] + 2\frac{X_1[n]-X_2[n]}{\pi}\sin\theta\right)^2\right] d\theta \\
&+ \frac{1}{4}\frac{2}{\pi}\int_{\pi}^{3\pi/2}\left[\left(X_1[n] - 2\frac{X_1[n]+X_2[n]}{\pi}\cos\theta\right)^2 + \left(X_2[n] - 2\frac{X_1[n]+X_2[n]}{\pi}\sin\theta\right)^2\right] d\theta \\
&+ \frac{1}{4}\frac{2}{\pi}\int_{3\pi/2}^{2\pi}\left[\left(X_1[n] + 2\frac{X_1[n]+X_2[n]}{\pi}\cos\theta\right)^2 + \left(X_2[n] + 2\frac{X_1[n]+X_2[n]}{\pi}\sin\theta\right)^2\right] d\theta
\end{aligned} \tag{6.56}
$$

$$
= 0.59(X_1[n]^2 + X_2[n]^2) \tag{6.57}
$$

The integral above is evaluated numerically.

Thus, the state magnitude decreases by a factor of $\frac{1}{0.59} = 1.69$ in expectation with the optimal control (for the given side information, i.e. quadrant for θ_n). This is independent of the time n. Hence, the tolerable mean-squared growth rate is $a_2^* = \sqrt{1.69} = 1.29$. ■

Lemma 6.6.5. *The logarithmic decay rate with two bits of side information, $Z[n] \in \{Q_1, Q_2, Q_3, Q_4\}$, for the system (6.47) is at least .47, and the tolerable growth rate is thus at least $a = 1.61$.*

Proof: Our aim is to minimize the terminal cost

$$
J_N(\mathbf{X}[N+1]) = \mathbb{E}\left[\log(X_1[N]^2 + X_2[N]^2)\right]. \tag{6.58}
$$

For every realization of the side information $Z[n]$ we choose to apply $U[n]$ that minimizes the mean-squared state magnitude. This is not the optimal strategy, but it provides a bound on the decay that is possible when an optimal strategy for the log-moment might be applied. Then, we can numerically evaluate the cost integral as

$$
\begin{aligned}
&J_n^*(\mathbf{X}[n], Z[n]) \\
&= \frac{1}{4}\frac{2}{\pi}\int_0^{\pi/2} \frac{1}{2}\log\left[\left(X_1[n] - 2\frac{X_1[n]+X_2[n]}{\pi}\cos\theta\right)^2 + \left(X_2[n] - 2\frac{X_1[n]+X_2[n]}{\pi}\sin\theta\right)^2\right] d\theta \\
&+ \frac{1}{4}\frac{2}{\pi}\int_{\pi/2}^{\pi} \frac{1}{2}\log\left[\left(X_1[n] + 2\frac{X_1[n]-X_2[n]}{\pi}\cos\theta\right)^2 + \left(X_2[n] + 2\frac{X_1[n]-X_2[n]}{\pi}\sin\theta\right)^2\right] d\theta \\
&+ \frac{1}{4}\frac{2}{\pi}\int_{\pi}^{3\pi/2} \frac{1}{2}\log\left[\left(X_1[n] - 2\frac{X_1[n]+X_2[n]}{\pi}\cos\theta\right)^2 + \left(X_2[n] - 2\frac{X_1[n]+X_2[n]}{\pi}\sin\theta\right)^2\right] d\theta \\
&+ \frac{1}{4}\frac{2}{\pi}\int_{3\pi/2}^{2\pi} \frac{1}{2}\log\left[\left(X_1[n] + 2\frac{X_1[n]+X_2[n]}{\pi}\cos\theta\right)^2 + \left(X_2[n] + 2\frac{X_1[n]+X_2[n]}{\pi}\sin\theta\right)^2\right] d\theta
\end{aligned} \tag{6.59}
$$

$$
= (-.477) + \log\left(X_1[n]^2 + X_2[n]^2\right). \tag{6.60}
$$

Hence, for this system with two bits of side information the average logarithmic decay is at least .477 and the tolerable growth $a_0^* \geq e^{.477} = 1.61$. ■

6.6.3 A spinning observer

The behavior of the corresponding estimation problem presents a sharp contrast to the control problem. Consider the system below, which seems to be a dual to the control problem earlier. There is no randomness in the system evolution and the target location scales at rate a. However, the objective is now to estimate the position of the target $\begin{bmatrix} X_1[n+1] & X_2[n+1] \end{bmatrix}^T$, using the observations $Y[n]$. $Y[n]$ is obtained by projecting the position of the target along the direction θ_n which is chosen uniformly at random as in the control case. (See Fig. 6.10)

$$
\begin{aligned}
\begin{bmatrix} X_1[n+1] \\ X_2[n+1] \end{bmatrix} &= a \begin{bmatrix} X_1[n] \\ X_2[n] \end{bmatrix}, \\
Y[n] &= \begin{bmatrix} \cos\theta_n & \sin\theta_n \end{bmatrix} \begin{bmatrix} X_1[n] \\ X_2[n] \end{bmatrix}.
\end{aligned} \tag{6.61}
$$

Complete information, θ_n is perfectly known at time n

Let us consider the case where θ_n is perfectly known to the observer.

Theorem 6.6.6. *The average logarithmic decay rate for estimation error the system (6.61) is ∞ if θ_n is perfectly known to the controller.*

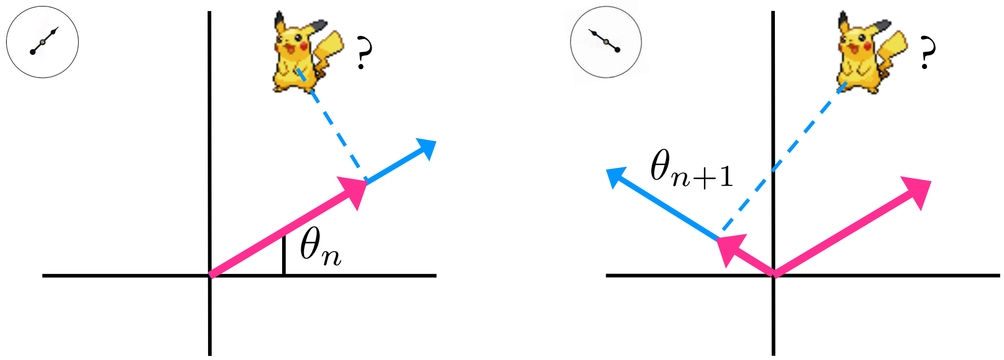


Figure 6.10: The objective of the observation problem is to estimate the position of the yellow Pikachu target at time n using the observation $Y[n]$. The magnitude of the projection in a direction (i.e. the pink vector) is revealed to the observer at each time. Combining the observations at time n and $n+1$ the observer can have a perfect estimate in two time steps.

Proof: At time $n+1$ the controller has access to:

$$\begin{bmatrix} Y[n] \\ Y[n+1] \end{bmatrix} = \begin{bmatrix} \cos\theta_n & \sin\theta_n \\ a\cos\theta_{n+1} & a\sin\theta_{n+1} \end{bmatrix} \begin{bmatrix} X_1[n] \\ X_2[n] \end{bmatrix}. \tag{6.62}$$

The observation from time n is retained by the observer in memory, and at time $n+1$ the observer gets a second set of equations regarding the position of the target.

Since θ_n and θ_{n+1} are chosen uniformly at random, with probability 1 these equations will be linearly independent. $\begin{bmatrix} \cos\theta_n & \sin\theta_n \\ a\cos\theta_{n+1} & a\sin\theta_{n+1} \end{bmatrix}$ is invertible. Hence, the estimation error at the end of two time steps will be zero. In sharp contrast to the control case, this gives an estimation error decay rate of infinity!

■

In this sense, the control problem considered earlier is “harder” than the observation problem. There, the maximal decay rate in both the mean-squared and logarithmic cases is finite. It turns out this has to do with the fact that in the observation problem, the memory of the past observation direction is useful to reduce estimation error at the current time. Information about future control directions plays a similar role to help the control problem as we will see in Sec. 7.4.1.

It turns out that partial information about θ_n plays very different roles in the observation and in the control problem.

Remark 6.6.7. Surprisingly, the two-step observability result is quite fragile. We know from the arguments in Chapter 4 that even a slight continuous uncertainty regarding θ_n renders the estimation problem impossible if $|a| > 1$.

6.7 Appendix

6.7.1 Logarithmic stabilizability

Theorem 6.7.1. *A system is **logarithmically stabilizable** as defined in Def. 5.2.2 if and only if there exists a strategy U_0^∞, with $U[i] < \infty$ for all $0 \leq i \leq \infty$ such that* $\liminf_{n\to\infty} \sum_{i=0}^{n-1} \mathbb{E}\left[G_{\mathcal{S}}(X[i], U[i], F_{\mathcal{T}[i]})\right] > -\infty$.

Proof: We assume $\mathbb{E}[\log ||X[0]||]$ exists and is finite.

($\Longleftarrow$)

$$\sum_{i=0}^{n-1} \mathbb{E}\left[G_{\mathcal{S}}(X[i], U[i], F_{\mathcal{T}[i]})\right] = \mathbb{E}\left[-\log \frac{||X[n+1]||}{||X[0]||}\right] \tag{6.63}$$

$$= \mathbb{E}[\log ||X[0]||] - \mathbb{E}[\log ||X[n+1]||] \tag{6.64}$$

Hence, there exists $N < \infty, K > -\infty$ such that

$$\mathbb{E}[\log ||X[0]||] - \mathbb{E}[\log ||X[n+1]||] > K \ \forall\ n > N \tag{6.65}$$

$$\Longrightarrow \mathbb{E}[\log ||X[n+1]||] < -K + \mathbb{E}[\log ||X[0]||] \ \forall\ n > N \tag{6.66}$$

This implies that $\mathbb{E}[\log ||X[n]||]$ is bounded for all $n > N+1$. Now consider $\mathbb{E}[\log ||X[n]||]$ for $0 \leq n \leq N$. Since N is finite, $\mathbb{E}[\log ||X[0]||]$ exists and is finite, and $U[i] < \infty$, $\max\limits_{0\leq n\leq N} \mathbb{E}[\log ||X[n]||] < \infty$. Combining this with eq. (6.66), there exists $M < \infty$ such that $\mathbb{E}[\log ||X[n]||] < M$ for all n. Hence, the condition above implies Def. 5.2.2.

($\Longrightarrow$)

If there exists $M < \infty$ such that $\mathbb{E}[\log ||X[n]||] < M$ for all n then

$$\lim_{n\to\infty} \mathbb{E}[\log ||X[0]||] - \mathbb{E}[\log ||X[n+1]||] > \mathbb{E}[\log ||X[0]||] - M > -\infty \tag{6.67}$$

$\mathbb{E}[\log ||X[n]||] < M$ also implies that $U[i] < \infty$ for all $0 \leq i \leq \infty$.

Hence, Def. 5.2.2 implies the condition above. ■

We also define strong logarithmic stabilizability below, in a stricter sense than logarithmic stability in Def. 5.2.2.

Definition 6.7.2. A system is **strongly logarithmically stabilizable** if there exists a strategy U_0^∞ such that $\lim_{n\to\infty} \sum_{i=0}^{n-1} \mathbb{E}\left[G_{\mathcal{S}}(X[i], U[i], F_{\mathcal{T}[i]})\right] = \infty$.

Strong logarithmic stabilizability essentially requires the state to asymptotically decay to 0. This is similar to notions of stability defined in control before, such as asymptotic stabilizability [15]. Thm. 6.7.3 establishes that strong logarithmic stability is a stronger notion than logarithmic stability.

Theorem 6.7.3. *Strong logarithmic stabilizability (Def. 6.7.2) implies logarithmic stabilizability (Def. 5.2.2).*

Proof:

$$\sum_{i=0}^{n-1} \mathbb{E}\left[G_{\mathcal{S}}(X[i], U[i], F_{\mathcal{T}[i]})\right] = \mathbb{E}\left[-\log \frac{||X[n+1]||}{||X[0]||}\right] \tag{6.68}$$

$$= \mathbb{E}[\log ||X[0]||] - \mathbb{E}[\log ||X[n+1]||] \tag{6.69}$$

Thus, strong logarithmic stabilizability implies that $\mathbb{E}[\log ||X[n]||] \to -\infty$ as $n \to \infty$ ($\mathbb{E}[\log ||X[0]||]$ is a finite constant). Hence, there exists $M \in \mathbb{R}$, $M < \infty$ such that $\mathbb{E}[\log ||X[n]||] < M$ for all n.

■

Strong logarithmic stability is also related to logarithmic stability in the sense that if a linear system without additive noise is strongly logarithmically stabilizable, i.e. the system state $X[n]$ is going to zero exponentially fast in n, then the same linear system with appropriately bounded additive noise will be logarithmically stabilizable.
On the other hand, a system might be logarithmically stable, but might not be strongly logarithmically stabilizable. Consider, for example, the system $X[n] = 5$ for all n. This system is logarithmically stabilizable, but is not strongly logarithmically stabilizable.

Theorem 6.7.4. *This theorem connects the average decay rate to strong logarithmic stabilizability.*

1. *If the average logarithmic decay rate $\bar{G}_{\mathcal{S}}(X[0], U_0^\infty, F_{\mathcal{T}_0^\infty}) > 0$ for some U_0^∞ then the system is strongly logarithmically stabilizable.*

2. *If the average logarithmic decay rate $\bar{G}_{\mathcal{S}}(X[0], U_0^\infty, F_{\mathcal{T}_0^\infty}) < 0$ for all U_0^∞ then the system is not strongly logarithmically stabilizable.*

Proof: (1) Suppose $\lim_{n\to\infty} \sum_{i=0}^{n-1} \mathbb{E}\left[G_{\mathcal{S}}(X[i], U[i], F_{\mathcal{T}[i]})\right] = K < \infty$. Then,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{E}\left[G_{\mathcal{S}}(X[i], U[i], F_{\mathcal{T}[i]})\right] = 0, \tag{6.70}$$

which is a contradiction. Hence, $\lim_{n\to\infty} \sum_{i=0}^{n-1} \mathbb{E}\left[G_{\mathcal{S}}(X[i], U[i], F_{\mathcal{T}[i]})\right] = \infty$ and the system is strongly logarithmically stabilizable.

(2)

$$
\begin{aligned}
&\bar{G}_{\mathcal{S}}(X[0], U_0^\infty, F_{\mathcal{T}_0^\infty}) < 0 && (6.71)\\
&\implies \lim_{n\to\infty} \frac{1}{n} \left(\mathbb{E}[\log||X[0]||] - \mathbb{E}[\log||X[n+1]||]\right) < 0 && (6.72)\\
&\implies \lim_{n\to\infty} \frac{1}{n} \left(-\mathbb{E}[\log||X[n+1]||]\right) < 0 \quad (\text{since } \lim_{n\to\infty} \frac{1}{n}\mathbb{E}[\log||X[0]||] = 0) && (6.73)\\
&\implies \lim_{n\to\infty} \frac{1}{n} \left(\mathbb{E}[\log||X[n+1]||]\right) > 0 && (6.74)\\
&\implies \lim_{n\to\infty} \mathbb{E}[\log||X[n+1]||] = \infty && (6.75)
\end{aligned}
$$

Thus,

$$
\begin{aligned}
&\lim_{n\to\infty} \sum_{i=0}^{n-1} \mathbb{E}\left[G_{\mathcal{S}}(X[i], U[i], F_{\mathcal{T}[i]})\right] && (6.76)\\
&= \lim_{n\to\infty} \sum_{i=0}^{n-1} \left(\mathbb{E}[\log||X[0]||] - \mathbb{E}[\log||X[n+1]||]\right) && (6.77)\\
&= \mathbb{E}[\log||X[0]||] - \infty && (6.78)\\
&= -\infty && (6.79)
\end{aligned}
$$

Since this is true for all U_0^∞ the system is not strongly logarithmically stabilizable.

■

In the example system (5.1) with $B[n] \sim \text{Uniform}[b_1, b_2]$, and $U[n] = -\frac{2}{b_1+b_2}X[n]$ for each n, $\bar{G}_{\mathcal{S}}(X[n], U_0^\infty, F_B) = \log|\frac{1}{a}| + \mathbb{E}\log|\frac{b_1+b_2}{b_1+b_2-2B[n]}|$. Hence if $\log|\frac{1}{a}| + \mathbb{E}\log|\frac{b_1+b_2}{b_1+b_2-2B[n]}| > 0$, the system is strongly logarithmically stabilizable using $U[n]$.

6.7.2 η-th moment stability

Theorem 6.5.9. $\lim_{\eta\to 0} a_\eta(U_0^\infty) = a_0(U_0^\infty)$

Proof: Consider

$$
\begin{aligned}
&\lim_{\eta\to 0}\sum_{i=0}^{n-1} -\frac{1}{\eta}\log \mathbb{E}\left[\frac{||X[i+1]||^{\eta}}{||X[i]||^{\eta}}\right] && (6.80)\\
=&\lim_{\eta\to 0}\sum_{i=0}^{n-1} -\mathbb{E}\left[\frac{||X[i+1]||^{\eta}}{||X[i]||^{\eta}}\log\left(\frac{||X[i+1]||}{||X[i]||}\right)\right]\frac{1}{\mathbb{E}\left[\frac{||X[i+1]||^{\eta}}{||X[i]||^{\eta}}\right]} && (6.81)\\
&\text{(take derivatives w.r.t } \eta \text{ and apply L'Höspital's rule)}\\
=&\sum_{i=0}^{n-1} -\mathbb{E}\left[1\cdot\log\left(\frac{||X[i+1]||}{||X[i]||}\right)\right]\frac{1}{1} && (6.82)\\
&\text{(taking the limit)}\\
=&\sum_{i=0}^{n-1} -\mathbb{E}\left[\log\left(\frac{||X[i+1]||}{||X[i]||}\right)\right] && (6.83)\\
=&\mathbb{E}\left[-\log\left(\frac{||X[n]||}{||X[0]||}\right)\right] && (6.84)\\
& && (6.85)
\end{aligned}
$$

Now consider,

$$
\begin{aligned}
&\lim_{\eta\to 0} a_{\eta}(U_0^{\infty}) && (6.86)\\
&= \lim_{\eta\to 0}\exp\left\{\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1} -\frac{1}{\eta}\log\mathbb{E}\left[\frac{||X[n+1]||^{\eta}}{||X[n]||^{\eta}}\right]\right\} && (6.87)\\
&= \lim_{n\to\infty}\exp\left\{\frac{1}{n}\lim_{\eta\to 0}\sum_{i=0}^{n-1} -\frac{1}{\eta}\log\mathbb{E}\left[\frac{||X[n+1]||^{\eta}}{||X[n]||^{\eta}}\right]\right\} && (6.88)\\
&= \lim_{n\to\infty}\exp\left\{\frac{1}{n}\mathbb{E}\left[-\log\left(\frac{||X[n]||}{||X[0]||}\right)\right]\right\} && (6.89)\\
&= a(U_0^{\infty}) && (6.90)
\end{aligned}
$$

■

Chapter 7

Future information in control

7.1 Introduction

This "forward-looking" chapter is about the value of information from the future. To be precise, we consider two examples of vector control systems where information about the future states of the actuation channel are available to the controller non-causally.

Before we discuss these examples, let us think about non-causal side information in communication. The capacity of an erasure channel with drop probability p_e is given by $1 - p_e$. This is because, on average, a $1-p_e$ fraction of the transmitted bits will get through the channel. Knowing when the erasures will happen might help the encoder schedule transmissions and save power. However, without any power limitations non-causal side information about the future erasure channel states cannot increase performance from a capacity perspective.

What about a memoryless AWGN channel? Knowing the channel noise perfectly here is certainly useful, however, receiving this perfect side information about the noise realization at time $t_1 + 1$ at time t_1, is exactly as useful as receiving it at time $t_1 + 1$ itself. For this perfect side information case, advance knowledge does not provide any advantage. However, in the case of imperfect side information, non-causal side information can be useful, as we know from the dirty paper coding results [22, 39]. Of course, non-causal side information can also be useful in the cases of channels with state and memory.

This chapter considers non-causal side information in vector control systems. In some problems, non-causal side information about unreliable actuation channels can help get around structural limitations in the system and significantly increase the control capacity of the system. The value of information and the ideas of preview control have been examined in control, as discussed in the introduction to Chapter 6. For example, Davis observes in [26], that future side information about noise realizations can effectively reduce a stochastic control problem to a set of deterministic ones: one for each realization of the noise sequence. The area of non-anticipative control characterizes the price of not knowing the future realizations of the noise [95, 35]. There has been a series of works that studies preview control. Preview information about disturbances can improve control performance [53]. Other related works and references for this also include [107, 74, 20]. Martins et al. [71] studied the impact of a preview, or non-causal side information (regarding an additive disturbance) on the frequency domain sensitivity function of the system.

The idea of lookahead in control is also reminiscent of the results in Gupta et al. [51], where precomputed future control values and actuator memory allows the system to be robust to more dropped control values. However, they rely on the key assumption that the controller can use an encoder to stack up control actions and the actuator has the ability to decode these, know what time it is, remember what it was sent in the past, and then apply the right control. This reduces the actuation channel to being a communication channel, which gives the result. The authors mention that real-world systems often have correlated losses, and it is reasonable to think we might have some knowledge of the future channel states that can help us plan controls even in the case where there is no memory in the actuator itself.

Future information has also been considered in other related fields. For example, Asnani and Weissman [4] consider the utility of finite lookahead for source coding through a dynamic programming framework, and Spencer et al. [103] look at the gains from knowing the future arrival sequence in queueing problems. The work here has a similar spirit.

The results and observations in this chapter are motivated by two problems discussed in the introduction. The intermittent Kalman filtering problem [102, 86], and the dual problem of control with dropped packets considered by Elia [30, 101], are examples of problems with multiplicative sensing and actuation channels respectively. The critical drop probability for mean-squared stable estimation of a system [86], with generic matrix A, is $\frac{1}{\lambda_{max}(A)^2}$. Thus the tolerable growth for mean-squared observability is dictated by the maximal eigenvalue. On the other hand, for the dual control problem, the critical erasure probability is $\frac{1}{\Pi\lambda(A)^2}$, and thus the tolerable growth for mean-squared stability is much lower. Why?

An obvious answer is that the control problem is structurally different, and that there is something about the physical aspect of control that makes the problem more difficult. However, in this chapter we show that the gap can also be thought of as an informational one. Non-causal side information to the controller about the erasure actuation channel can increase the tolerable growth rate of the control problem to match that of the corresponding observation problem. However, the same non-causal lookahead would not help the observation problem do any better.

The last part of the chapter reconsiders the spinning example from Chapter 6, this time in the context of non-causal side information. The continuous uncertainty on the actuation channel allows us to modulate the side information on a finer scale than with a binary erasure channel. The example more clearly illustrates why the discrepancy between the observation and control counterparts disappears with information from the future.

7.2 The gap between control and estimation

First, consider the vector system with a single scalar control input $U[n]$ sent over a real-erasure channel as below:

$$\vec{X}[n+1] = A \cdot \vec{X}[n] + \beta[n] \cdot B \cdot U[n],$$
$$\vec{Y}[n] = \vec{X}[n], \tag{7.1}$$

where $\vec{X}[n]$ is the $m \times 1$ state vector, A is an $m \times m$ constant matrix, B is an $m \times 1$ constant matrix, and $\beta[n]$ are i.i.d. Bernoulli$(1 - p_e)$ random variables.

Theorem 7.2.1. *(Elia [30], Schenato et al. [101]) The system eq. (7.3) is mean-square stabilizable if and only if $p_e < \frac{1}{\Pi\lambda(A)^2}$, where $\lambda(A)$ are the unstable eigenvalues of the matrix A, and (A, B) form a controllable pair.*

Remark 7.2.2. This result follows from the combination of [30] and [101], and we only give a high level picture here. Some of the general ideas here were also discussed in the Introduction, Chapter 1.

Since the multiplicative noise on the actuation channel made the problem difficult, Elia considered a simpler problem and restricted his exploration to stabilization using LTI controllers. For LTI controllers, he showed that $p_e < \frac{1}{\Pi\lambda(A)^2}$, where $\lambda(A)$ are the unstable eigenvalues of the matrix A, is the critical threshold probability for stabilization.

The proof for the necessary condition used a novel idea. It decouples the problem into two sub-problems: (1) the *stability* of a system with a multiplicative noise in the system dynamics loop and no control action allowed, and (2) the *stabilizability* of a system with an additive noise using a control action. The first problem is used to get a relationship between the properties of the multiplicative noise and the system transfer function. The second problem involves an implicit rate constraint induced by an additive noise channel. Ideas from rate-limited control (e.g. [110]) are used to give the $\frac{1}{\Pi\lambda(A)^2}$ bound.

[101, 55] used dyanamic programming techniques to show that in the case where packet losses over the actuation channel are acknowledged (i.e. the control transmission is over a "TCP-like" protocol (Fig. 1.6 from the introduction), separation between estimation and control holds. Further, these works provide LMI characterizations for the erasure probabilities and show that LTI control strategies are in fact optimal in the infinite horizon. The optimal control strategy is a linear function of the state-estimate, and the optimal controller gain converges to a constant in the infinite horizon case [101]. Though the paper focuses on time-varying cost functions, the technique used can show that linear strategies are optimal even for a finite time horizon if the state and control costs are time-invariant. The proof uses dynamic-programming-style backward induction on the state-covariance. Now, since the assumption of perfect observations in [30] is equivalent to the assumption of actuation channel acknowledgements, we can conclude that there is no loss of optimality in restricting to LTI systems and the bound from [30] holds generally.

For the special case where B is invertible, and U is a vector input, the maximum eigenvalue of the systems was identified as the bottleneck [55]. ■

As we discussed in the introduction, the $\frac{1}{\Pi\lambda(A)^2}$ threshold is significantly more stringent than the threshold for the dual intermittent Kalman filtering problem in eq. (7.2). We know now from [86] that for the generic case with A matrices with no cyclic structure in their eigenvalues, $\frac{1}{\lambda_{max}(A)^2}$ is the mean-squared observability threshold on the erasure probability.

$$\begin{aligned} \vec{X}[n+1] &= A \cdot \vec{X}[n] \\ \vec{Y}[n] &= \beta[n] \cdot \vec{X}[n] \end{aligned} \tag{7.2}$$

However, [86] also showed that this threshold becomes more stringent in the case where there is a periodicity to the eigenvalues. The observer can decode the state as soon as it receives enough equations: at least n equations are required to decode an n dimensional state. However, periodic eigenvalues could lead to redundant observations that are "aligned," i.e. they are scalar multiples of each other. These redundant observations do not help in decoding the state, which leads to the tighter threshold.

7.3 Understanding the gap

Now, we try to understand the informational differences between the observation and the control problems in the generic case. We consider a 2×2 control system

$$\begin{aligned} \begin{bmatrix} X_1[n+1] \\ X_2[n+1] \end{bmatrix} &= \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} X_1[n] \\ X_2[n] \end{bmatrix} + \beta[n] \begin{bmatrix} 1 \\ 1 \end{bmatrix} U[n], \\ \vec{Y}[n] &= \begin{bmatrix} X_1[n] \\ X_2[n] \end{bmatrix}. \end{aligned} \tag{7.3}$$

and the corresponding observation problem:

$$\begin{aligned} \begin{bmatrix} X_1[n+1] \\ X_2[n+1] \end{bmatrix} &= \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} X_1[n] \\ X_2[n] \end{bmatrix} \\ Y[n] &= \beta[n] \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} X_1[n] \\ X_2[n] \end{bmatrix}. \end{aligned} \tag{7.4}$$

Let $\lambda_{max} = |\lambda_1| > |\lambda_2| \geq 1$ be the system eigenvalues. $\beta[n]$ is a Bernoulli$-(1-p_e)$ random variable. For the control problem, the actuation channel transmission is successful when $\beta[n] = 1$. For the observation problem, a sensor observation is received when $\beta[n] = 1$.

First we note that the observer can decode the state perfectly as soon as *any* two observations are received. Since the arrivals form a Geometric sequence with parameter $(1-p_e)$, and the arrivals are independent of each other, effectively the success of the observation strategy is distributed as a Negative Bionomial random variable with parameters $(2, 1-p_e)$. However, the tail behavior of this is the same as that of a Geometric$-(1-p_e)$. On the other hand, to be able to set the state to zero, the controller needs to plan. *Any* two successful controls are not enough. If a pair of controls at times t_1 and t_2 are to be successful, the controller must know the value of t_2 in advance at time t_1. While for the observation problem, memory of past observation arrival times suffices, it does not for the control problem. Instead, what matters is the future.

Consider the behavior of a naive "planning" control strategy, which we call align-and-kill, in the 2×2 setting. At each time t_1, align-and-kill assumes that both the control action at time t_1 and the action at time t_1+1 will be successful, i.e. $\beta[t_1] = \beta[t_1+1] = 1$, and computes the control $U[t_1]$ accordingly. If the state is already aligned to the control direction available at time t_1, $U[t_1]$ sets the state to zero. If not, $U[t_1]$ does the job of "aligning" the state to the control direction at time $t_1 + 1$, so that it can be set to zero at $t_1 + 1$, if $\beta[t_1 + 1]$ does indeed turn out to be 1. Since this strategy needs two successes in a row, the waiting time for this strategy to be successful is behaves likes a Geometric$-(1-p_e)^2$ (in the tail sense), which is harsher than the observation problem. We will see later that align-and-kill seems to be asymptotically optimal in the regime of large eigenvalues.

The controller could get around this bottleneck if it had access to a vector control, i.e. two components $U_1[\cdot]$ and $U_2[\cdot]$ that were able to independently control both components X_1 and X_2. A controller could then set the state to zero with one successful action. Then, with this *structural* change, the waiting time for a successful control would be a geometric with parameter $(1 - p_e)$.

7.3.1 Infinite future side information

This section tries to capture the essence of the informational bottleneck on the controller. For this, consider a controller with infinite side information about the future. If the controller knew the entire sequence of $\beta[n]$, at each time t_1 it would also know the time of the next control success, t_2. Then it could easily set up the system so as to kill the state at time t_2. This idea is captured in the Theorem. 7.3.1.

Theorem 7.3.1. *The critical probability for the control problem eq. (7.3), with infinite lookahead for the sequence $\beta[n]$, is $\frac{1}{\lambda_{max}(A)^2}$.*

Proof: The proof follows from a time-reversal argument: the control problem is essentially the same as the observation problem if time is reversed and we are only waiting on the first two arrivals. The result builds on the arguments in [83] and the proof techniques used there. This depends on a few key facts. The first is that the observability Grammian and the controllability Grammian for the two systems being considered are transposes of each other. To avoid the complication of dealing with the full history, [83] focuses on using a suboptimal linear estimator in the ML-style. Basically, it simply inverts the observability Grammian once enough observations have been collected. This is suboptimal, but [83] shows that it is good enough to reach the $\frac{1}{\lambda_{max}(A)^2}$ threshold. Because the proof in [83] just used properties of the inverse of the observability Grammian, we can apply the same results to the controllability side and compute a "zero-forcing" control by inverting the relevant controllability Grammian. We just need to know the future actutation times so the appropriate Grammian can be inverted. ■

Remark 7.3.2. What this shows is that knowledge of the future channel behavior can get around the structural bottleneck posed by only having a scalar control to stabilize a vector system.

7.3.2 Fully periodic eigenvalues

Conditions for estimation and control with multiplicative noise can match in certain corner cases[1]. We look at the special case of periodic eigenvalues to get a better understanding of the planning bottleneck that the controller faces. Essentially, the structure in the system below removes the controllers ability to plan. This same structure causes redundant observations in the Kalman filtering problem, making this the "hardest" case.

Consider an n dimensional system with cyclic eigenvalues for A that also have period n. For this case, we can give a simple argument to show that the critical probability for the observation problem and the control problem are exactly the same: $\left|\frac{1}{\Pi\lambda(A)^2}\right|$. The relevant uncertainty growth in the control case is given by the product of the eigenvalues regardless of cycles in A. This "hardest" case in the estimation problem [86], seems to be the generic case for the dual control problem. In this case, with fully-periodic eigenvalues, future information would not be useful.

Consider eq. (7.3) with $\lambda_1 = 2$ and $\lambda_2 = -2$. These eigenvalues form a cycle of period two. We can further separate out the angle and the magnitude of the eigenvalues and write the gain matrix $A = 2\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}$. $\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}$ is a matrix that reflects the second component, and $\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}^2 = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$ is the identity matrix.

Theorem 7.3.3. *A greedy control strategy that projects the state vector in the orthogonal direction to* $B = \begin{bmatrix}1\\1\end{bmatrix}$, *namely* $U[n] = -\begin{bmatrix}1 & -1\end{bmatrix}\begin{bmatrix}X_1[n]\\X_2[n]\end{bmatrix}$, *is the optimal strategy for the system in eq.* (7.3), *with* $\lambda_1 = 2$ *and* $\lambda_2 = -2$.

Proof: The basic idea of the proof is that controls applied at even and odd times cannot substitute for each other. One of each is essential.

[1] Most of the results in this and in following sections of this chapter are joint work with my undergrad student Govind Ramnarayan [92].

Let the first arrival be at time n_1. Choose $u[n_1] = -X_1[n_1] + X_2[n_1]$. With this control, the projection of the state onto the $\begin{bmatrix}1\\1\end{bmatrix}$ direction is zeroed out.

Then, if the control is successful,

$$\begin{bmatrix}X_1[n_1+1]\\X_2[n_1+1]\end{bmatrix} = \begin{bmatrix}X_1[n_1+1]+X_2[n_1+1]\\-X_1[n_1+1]-X_2[n_1+1]\end{bmatrix}, \tag{7.5}$$

is orthogonal to $\begin{bmatrix}1\\1\end{bmatrix}$. This implies that $\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}^{2r+1}\begin{bmatrix}X_1[n_1+1]\\X_2[n_1+1]\end{bmatrix}$ is aligned to $\begin{bmatrix}1 & 1\end{bmatrix}^T$ for every r. Consequently, the state can be set to zero at any arrival at any "odd" timestep $n_1 + 2r + 1$ for any r.

All time steps of the form $n_1 + 2r$, which we call "even" timesteps, would have the state vector orthogonal to $\begin{bmatrix}1 & 1\end{bmatrix}^T$ rendering the control useless. Thus, we need two arrivals that are spaced an odd number of steps(i.e. $2r + 1$ steps) apart to set the state to zero. Hence, the strategy proposed must be optimal. ■

These arguments for the 2×2 case generalize to systems of dimension-k with an eigenvalue cycle of period-k and the same results hold: the estimation critical probability and the control critical probability are exactly the same. Furthermore, non-causal knowledge of the pattern of arrivals would have no impact on the rate of the decay of the system state. Here, side information has no value because there is essentially no flexibility to plan, since there are no substitutes between timeslots (mod k).

7.3.3 Partial future side information

This section explores the behavior of a family of systems using different control strategies and various amounts of future information. While full non-causal information about the channel behavior might be unrealistic, it is often possible to have side information about the immediate future of the channel. Here we implement a dynamic program to calculate the value of knowing just the value of the actuation channel at the next time step. This is measured through the critical drop probability of the system: which channels cause the system to go unstable? This inverse of the critical drop probability can act as a proxy for the tolerable growth of the system. The details of this computation are given in Appendix 7.5.1.

We considered a family of systems (eq. (7.3)) with the eigenvalues scaling together such that $\frac{\lambda_1}{\lambda_2} = 1.18$. The x-axis in Fig. 7.1 represents the maximal eigenvalue of the system on log scale and the y-axis plots the critical drop probability for that system on log scale. For the problem considered in [30], with no future information (lookahead), the critical probability follows a line of slope -4, i.e. $\frac{1}{\lambda_1^2\lambda_2^2}$. This is the green line. For the problem with full future lookahead, the critical probability is given by $\frac{1}{\lambda_1^2}$, which is the topmost yellow line (slope -2).

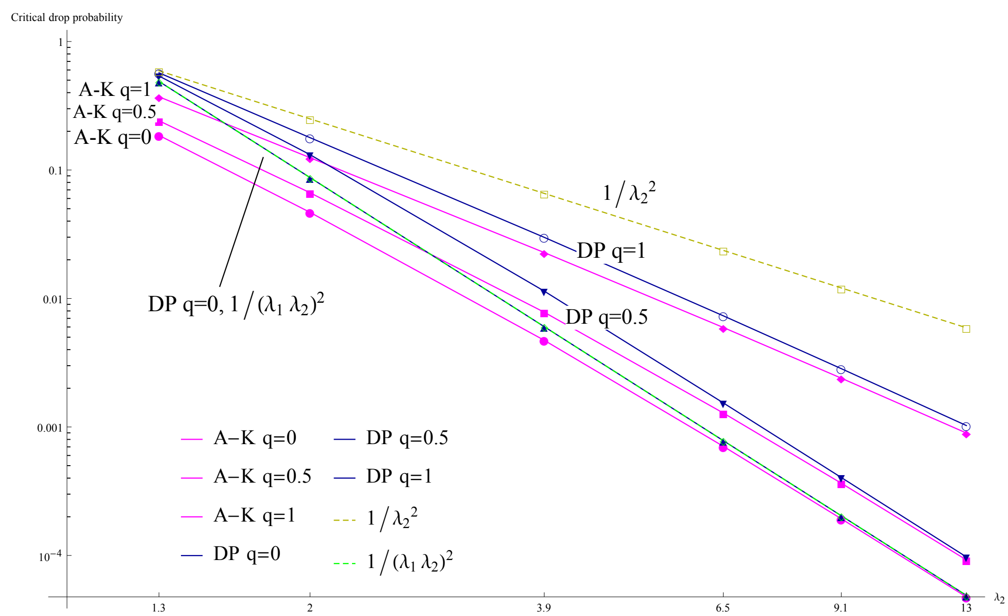


Figure 7.1: Critical erasure probability versus magnitude of the maximum eigenvalue $\frac{\lambda_1}{\lambda_2} = 1.18$.

As expected optimal dynamic programming solution with no lookahead decays exactly as the optimal strategy: the lines for $\frac{1}{(\lambda_1\lambda_2)^2}$ and the optimal dynamic programming strategy ($q = 0$) are on top of each other [55].

Next we consider systems with finite amounts of future lookahead. The parameter q in the figure represents the probability with which lookahead is received. Let us consider the case $q = 1$. This means that at each time t_1, the controller is told whether the control action at time t_1+1 will be dropped or not. These bits of side information can be stored in memory, effectively the controller knows the channel behavior at both time t_1 and $t_1 + 1$ at time t_1. While ideally the controller would like to know the time of the next successful action t_2, can information about $t_1 + 1$ help? In the best case scenario of course, the controller learns that the at time $t_1 + 1$ is going to be successful, and computes the control for the current time t_1 accordingly. However, the information that $t_1 + 1$ will be dropped reduces its uncertainty about the next arrival t_2. In particular, it knows that there is no point in planning for a successful control at time $t_1 + 1$. The improvement due to this, i.e. the increase in "control capacity," is seen with the blue line labelled $q = 1$, with slope slightly greater than -3.

We can also compute the value of unreliable one-step lookahead, where information about t_1+1 is received with probability $q = 0.5$. This calculation gives the second blue line, labelled $q = 0.5$. This also gives a scaling improvement on the control capacity, though not as drastic. We can also see that as the actuation channel gets more unreliable the value of the extra bit of lookahead increases. From a practical perspective, a control designer might prefer a more predictable channel, even though it might be more unreliable on average. That the slopes are different suggests that the side information has a *scaling* effect on the mapping between the eigenvalues and the critical erasure probability.

Finally, we also consider the behavior of the naive control strategy that we call align-and-kill. At each time t_1, align-and-kill assumes that both the control action at time t_1 and the action at time $t_1 + 1$ will be successful and computes controls accordingly. Align-and-kill with guaranteed lookahead ($q = 1$) about $t_1 + 1$ behaves as usual if there will be a successful action at $t_1 + 1$ and if not plans for a "kill" at time $t_1 + 2$. Align-and-kill with unreliable lookahead ($q = 0.5$) hedges between the two cases. What is interesting is that the pink lines represented by the align-and-kill strategies seem to converge to the blue optimal control thresholds at high eigenvalues. This suggests that as the system growth gets more unstable, planning for the next timestep using an align-and-kill approach might be the best thing a controller can do, and this warrants further investigation.

The align-and-kill strategy is reminiscent of the zero-forcing equalization strategy in communication. Parallel to how zero-forcing is asymptotically optimal at high SNR, it seems the naive align-and-kill strategy is asymptotically optimal here. This supports the observations from Park's dissertation [83]: the magnitude of the eigenvalues of the control system seem to correspond somehow to SNR in wireless communication.

7.4 A toy example

This section finally comes back to the toy spinning system considered in the previous chapter. Here we can more clearly see how one step of lookahead about the future actuation channel state can instantly bridge the gap between the observation and the control problem. Using a continuous uncertainty on the actuation channel lets us have finer control on the side information we can provide.

7.4.1 Future information

The last chapter showed that the control problem with perfect information about θ_n at time n is "harder" than the observation problem with perfect information about θ_n. Here, we show that if the controller is given access to the value of the control direction at the future timestep as well as the control direction at the current time, the decay rate of the system is infinity, matching that of the observation problem!

Consider the system (6.35), where at time n the controller is given side information $Z[n] = \{\theta_n, \theta_{n+1}\}$. This allows the controller to compute, at time n, a pair of zero-forcing controls $U[n]$ and $U[n+1]$, such that the state is 0 at time $n + 2$. (See Fig. 7.2)

Theorem 7.4.1. *The average logarithmic decay rate with of the system (6.35) if both $Z[n] = \{\theta_n, \theta_{n+1}\}$ are known to the controller at time n is infinite.*

Proof: We consider the evolution of two time steps of the system. $\mathbf{X}[n + 2]$ in terms of $\mathbf{X}[n]$ is:

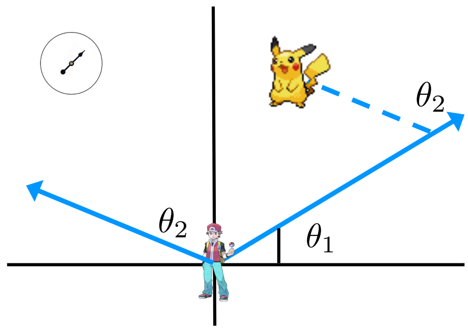


Figure 7.2: If the control direction at time 2 is known to the controller at time 1 it zero-force the state in 2 time steps as shown.

$$\begin{bmatrix} X_1[n+2] \\ X_2[n+2] \end{bmatrix} = a^2 \begin{bmatrix} X_1[n] \\ X_2[n] \end{bmatrix} + \begin{bmatrix} a^2 \cos\theta_n & a\cos\theta_{n+1} \\ a^2 \sin\theta_n & a\cos\theta_{n+1} \end{bmatrix} \begin{bmatrix} U[n] \\ U[n+1] \end{bmatrix} \tag{7.6}$$

Thus, if both θ_n and θ_{n+1} are known at time n the controller can solve for $U[n], U[n+1]$ such that $\mathbf{X}[n+2] = \vec{0}$.

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = a^2 \begin{bmatrix} X_1[n] \\ X_2[n] \end{bmatrix} + \begin{bmatrix} a^2 \cos\theta_n & a\cos\theta_{n+1} \\ a^2 \sin\theta_n & a\cos\theta_{n+1} \end{bmatrix} \begin{bmatrix} U[n] \\ U[n+1] \end{bmatrix} \tag{7.7}$$

Since the state goes to zero in a finite number of steps the average logarithmic decay rate is infinite!

■

In a sense, the value of the side information θ_{n+1} at time n was infinite. This future lookahead seems to bridge the gap between the control problem and the observation problem. For the observation problem, knowledge of the current and past direction led to an infinite decay rate. For the control problem, it seems that knowledge of the current and future direction gives us this.

The natural question then is: are just a few bits about θ_{n+1} valuable to the controller? Does partial information improve the control performance? (In the observation problem, we note that partial side information (about the current or past direction) was just as good as no side information at all.)

Partial future information

We explore the same system as (6.35)

$$\begin{bmatrix} X_1[n+1] \\ X_2[n+1] \end{bmatrix} = a \begin{bmatrix} \cos\Phi_n & \sin\Phi_n \\ -\sin\Phi_n & \cos\Phi_n \end{bmatrix} \left(\begin{bmatrix} X_1[n] \\ X_2[n] \end{bmatrix} + U[n] \begin{bmatrix} \cos\Theta_n \\ \sin\Theta_n \end{bmatrix} \right),$$
$$Y[n] = \begin{bmatrix} X_1[n] \\ X_2[n] \end{bmatrix}. \tag{7.8}$$

except now the control action $U[n]$ is a function of $Y[n]$ and $Z[n] = \{\theta_n, Z_1[n]\}$, where now $Z_1[n]$ contains some bits of information about θ_{n+1}. $Z_1[n]$ contains quantization information about θ_{n+1} just as in Sec. 6.6.2.

"Useful" partial side information about the future direction has a slightly different behavior than partial information about the present. What is one bit of side information in this context? Let us consider the natural choice of dividing the space into two half-spaces, $\mathcal{H}_1 = [0, \pi)$ and $\mathcal{H}_2 = [\pi, 2\pi)$, as we did in Chapter 6. Revealing $Z_1[n] \in \mathcal{H}_1$ actually constitutes 0 bits of side information at time n, since it does not actually reduce the space of possible directions from the perspective of control. Since θ_{n+1} will be perfectly revealed at time $n+1$, and the controller is allowed to choose a negative control, knowing that $\theta_{n+1} \in [0, \pi)$ at time n does not reduce any relevant uncertainty about the future control action.

What about dividing the space as $Q_1 = [0, \frac{\pi}{2}), Q_2 = [\frac{\pi}{2}, \pi), Q_2 = [\pi, \frac{3\pi}{2})$ and $Q_4 = [\frac{3\pi}{2}, 2\pi)$? This finer division is the one that actually constitutes one-bit of side information. For this, we reveal side information $Z_1[n] = \psi_n$, which means that we know $\theta_{n+1} \in [\psi_n, \psi_n + \frac{\pi}{2}]$.

Since ϕ_n is random, the reorientation of the axis can be captured by thinking of ψ_n as uniformly distributed from $[0, 2\pi]$. This lets us maintain the target at the same position across time, but at the same time allow for the spinning of the axis[2].

We examine this problem using dynamic programming for a time-horizon of two. What is the advantage of having lookahead about θ_1 at time 0, when θ_0 is perfectly known at time 0? We examine this for $k = 0, 1, 2, 3$ bits of lookahead about the control direction below. Since we want to look at the advantage provided by lookahead, a natural comparison is a two-step comparison, i.e. to look at $-\frac{1}{2}\log\mathbb{E}[\frac{||X[2]||^2}{||X[0]||^2}]$. This is the decay in the 2nd moment of the system in two time steps. To have the same "units" to compare this to the numbers in the previous section, the table here uses $-\frac{1}{4}\log\mathbb{E}[\frac{||X[2]||^2}{||X[0]||^2}]$, to provide numbers to compare to a "one-step" decay rate. We are essentially pretending that the decay at time 0 and at time 1 is the same, which in reality in this setup is not the case[3].

[2]This spinning of the axis is essential to preserve the symmetry of the problem. If we maintained the same axis, but we only learned whether θ_{n+1} was in $[0, \pi)$ or $[\pi, 2\pi)$ each time, then the optimal control would never move the target along the Y-axis at all. The intermediate spin ϕ_n prevents this and essentially gives us information about a new partition of the space at each time.

[3]Other preliminary calculations indicate that $-\frac{1}{T}\log\mathbb{E}[\frac{||X[T]||^2}{||X[0]||^2}]$ converges as T grows and is the appro-

Theorem 7.4.2. *The one-step 2nd moment and logarithmic decay rate for $k = 0, 1, 2, 3$ with bits of lookahead about the control direction θ_{n+1} (and perfect information about θ_n at time n), time horizon $N = 2$ for system (6.35) are as given in Table. 7.1.*

Table 7.1: System growth as a function of non-causal side information

Future side-info		2nd moment tolerable growth, a_2^*	Logarithmic tolerable growth, a_0^* lower bound
0 bits		$\sqrt{2}$ =1.41	2
1 bit		1.50	2.28
2 bits		1.74	2.92
3 bits		2.01	3.93
∞ bits		∞	∞

Proof: We set up the dynamic program to give the mean-squared decay rate for one bit of side information regarding θ_{n+1} and perfect knowledge of θ_n at time[4] n.

Let N denote the time horizon. Our aim is to minimize the terminal cost

$$J_N(\mathbf{X}[N+1]) = \mathbb{E}\left[X_1[N]^2 + X_2[N]^2\right]. \tag{7.9}$$

At each time $n < N$, the controller receives the current direction θ_n, as well as one bit of lookahead about the next direction θ_{n+1}. This lookahead comes in the form of ψ_n, such that $\theta_{n+1} \in [\psi_n, \psi_n + \frac{\pi}{2}]$. Thus, the controller knows the quadrant of the next direction. At each time step, the system is spun around uniformly. ψ_n is uniformly distributed from $[0, 2\pi]$.

Then the cost-to-go function at time n by $J_n(X_1[n], X_2[n], \theta_n, \psi_n, u[n])$ with current position $(X_1[n], X_2[n])$, control direction θ_n, lookahead direction ψ_n and control amplitude U_n is

$$J_n^*(X_1[n], X_2[n], \theta_n, \psi_n) = \min_{u[n]} J_n(X_1[n], X_2[n], \theta_n, \psi_n, Un]). \tag{7.10}$$

Note that $\mathbf{X}[n+1] = \begin{bmatrix} X_1[n] - U[n]\cos\theta_n & X_2[n] - U[n]\sin\theta_n \end{bmatrix}^T$. Then for $0 \leq n < N$, the evolution of the cost-to-go is determined as follows:

$$J_n(X_1[n], X_2[n], \theta_n, \psi_n, U[n]) =$$
$$\frac{1}{2\pi}\int_0^{2\pi}\left\{\frac{2}{\pi}\int_{\psi_n}^{\psi_n+\frac{\pi}{2}} J_{n+1}^*(X_1[n] - U[n]\cos\theta_n, X_2[n] - U[n]\sin\theta_n, \theta, \psi)\ d\theta\right\}\ d\psi. \tag{7.11}$$

priate quantity to look at the calculate the mean-squared decay rate in cases of not just one-step but even larger lookahead.

[4]The calculations for $k = 2, 3$ bits of information and logarithmic decay are similar to the calculations in Sec. 6.6.2 and are omitted.

The inner integral over θ captures the lookahead: θ_{n+1} is realized between ψ_n and $\psi_n + \frac{\pi}{2}$. The outer integral captures the fact at time n there is no knowledge about the control direction at time $n+2$ and so the lookahead direction at time $n+1$, ψ_{n+1} is uniformly distributed over 0 to 2π.

This program was implemented in Mathematica to calculate the numbers in Table 7.1. The logarithmic growth uses the optimal strategy for the the 2-nd moment, but looks at the 0-th moment (log) decay rate, to calculate a lower bound on the logarithmic tolerable growth. ■

7.5 Appendix

7.5.1 Dynamic program

This appendix gives a sketch of the key steps in the dynamic program used to calculate the value of unreliable partial side information in Sec. 7.3.3 and Fig. 7.1, for the system in eq. (7.3). Let $B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. A is the system matrix. Let p be the probability of a packet being dropped at any given time. Let q be the probability that at time n we get lookahead about time $n+1$.

- Let $J_{a,n}(\vec{X}n])$ be the cost-to-go at time n if there is an actuator success at the current time n.
- Let $J_{d,n}(\vec{X}n])$ be the cost-to-go at time n if there is an actuator drop at the current time n.
- Let $J_{b,n}(\vec{X}n])$ be the cost-to-go at time n if the controller does not know if there is a success or a drop at the current time n.

We will use dynamic programming to calculate the evolution of these three costs separately.

1. Case 1: Arrival at time n.
   - At time n we also get lookahead probability q, and this indicates an arrival at $n+1$ with probability p. In this case, the cost-to-go $J_{a,n}(\vec{X}[n], u[n]) = J^*_{a,n+1}(A\vec{X}[n] + Bu[n])$. Here $\vec{X}[n+1] = A\vec{X}[n] + Bu[n]$ since we know the control went through at time n.
   - With probability $1-p$ however, the lookahead (which comes with prob q) indicates no arrival. Then, $J_{a,n}(\vec{X}[n]) = J^*_{d,n+1}(A\vec{X}[n] + Bu[n])$.
   - Finally, with probability $(1-q)$ we get no looakahead at all. In this case, $J_{a,n}(\vec{X}[n]) = J^*_{b,n+1}(A\vec{X}[n] + Bu[n])$.

So in total we have:

$$J_{a,n}(\vec{X}[n], u[n]) = qpJ^*_{a,n+1}(A\vec{X}[n] + Bu[n]) + q(1-p)J^*_{d,n+1}(A\vec{X}[n] + Bu[n]) + (1-q)J^*_{b,n+1}(A\vec{X}[n] + Bu[n]) \tag{7.12}$$

So,

$$J^*_{a,n}(\vec{X}[n]) = \min_{u[n]} qpJ^*_{a,n+1}(A\vec{X}[n] + Bu[n]) + q(1-p)J^*_{d,n+1}(A\vec{X}[n] + Bu[n]) + (1-q)J^*_{b,n+1}(A\vec{X}[n] + Bu[n]) \tag{7.13}$$

2. Case 2: No arrival at time n. In this case $u[n]$ doesn't play a role, since we know it is not going to get through. The structure of future information remains the same as in the case above, except with state $A\vec{X}[n]$.

$$J_{d,n}(\vec{X}[n], u[n]) = qpJ^*_{a,n+1}(A\vec{X}[n]) + q(1-p)J^*_{d,n+1}(A\vec{X}[n]) + (1-q)J^*_{b,n+1}(A\vec{X}[n]) \tag{7.14}$$

Since there is no $u[n]$ to optimize over

$$J^*_{d,n}(\vec{X}[n]) = qpJ^*_{a,n+1}(A\vec{X}[n]) + q(1-p)J^*_{d,n+1}(A\vec{X}[n]) + (1-q)J^*_{b,n+1}(A\vec{X}[n]) \tag{7.15}$$

3. Case 3: No information about the arrival at time n. In this case we hedge our bets over what would happen at time n.

$$J_{b,n}(\vec{X}[n], u[n]) = pJ_{a,n}(\vec{X}[n], u[n]) + (1-p)J_{d,n}(\vec{X}[n], u[n]) \tag{7.16}$$

Now, to calculate the optimal cost:

$$J^*_{b,n}(\vec{X}[n]) = \min_{u[n]} pJ_{a,n}(\vec{X}[n], u[n]) + (1-p)J_{d,n}(\vec{X}[n], u[n]) \tag{7.17}$$

$$= \min_{u[n]} p\Big(qpJ^*_{a,n+1}(A\vec{X}[n] + Bu[n]) + q(1-p)J^*_{d,n+1}(A\vec{X}[n] + Bu[n]) + (1-q)J^*_{b,n+1}(A\vec{X}[n] + Bu[n])\Big) + (1-p)\Big(qpJ^*_{a,n+1}(A\vec{X}[n]) + q(1-p)J^*_{d,n+1}(A\vec{X}[n]) + (1-q)J^*_{b,n+1}(A\vec{X}[n])\Big) \tag{7.18}$$

$$= pJ^*_{a,n}(\vec{X}[n]) + (1-p)J^*_{d,n}(\vec{X}[n]) \tag{7.19}$$

Standard dynamic programming techniques can be used from here.

Chapter 8

Conclusion

This dissertation presents an information-theoretic framework to understand the stability of active control systems, and takes one step forward towards bridging the fields of communication and control. The aesthetic provides an angle by which we can measure the impact of uncertainty in control using information-theory tools. While much of the past work has focused on understanding information bottlenecks due to unreliable sensing channels, this dissertation provides tools that help understand uncertainties on the actuation channel. Traditionally, control systems have been seen as *sources* of uncertainty, and recent work [49, 83] has highlighted the *signaling* and *channel* aspects of control systems. The results here on control capacity show that there is a *sink* nature to the information in control systems. Controllers can "communicate" (in the broad sense of reducing uncertainty about the world) by removing unwanted bits from the system.

All of the results in this dissertation were discovered by explorations using simplified or toy models. The carry-free models that extended from wireless network information theory helped port the understanding of signal interactions in networks to understand the interactions between a control action and the system state. These models influenced the definitions of control capacity and provided strategy guidelines for proof techniques. They also clearly illustrate how uncertainty in parameters and the resulting multiplicative noise can have fundamentally different behavior than additive noise. A different toy model, the spinning system, helped understand phenomena in vector systems that are new compared to scalar systems. Previous work used diagonalization to view vector control systems as collections of scalar non-interacting systems. In the case of unreliability introduced through actuation channels, diagonalization does not quite work because the approach cannot capture the need to coordinate and the related flexibility to plan across the interconnected scalar parts of the systems.

What is the grand vision of this work? Ideally, we would want to provide a unified perspective on uncertainty and unreliability in information theory, robust control and machine learning. As sensing data about our surroundings becomes ubiquitous and processing gets cheaper, structural control and decision-making will be the cause of limitations in engineering systems. Data analysis has ceased to be the pressure point. However, we do not yet have a good paradigm to tell us which data we should collect and analyze in the first place. A universal "information theory" that incorporates goals and semantics that can work in a decision-making-framework is desired.

8.1 Future work

While the work in Chapter 5 lays the foundation to understand control capacity, many ideas require a fuller development.

- It remains to be seen how these ideas will formally extend to more general non-linear and vector systems. What about systems that include both additive and multiplicative noise? The work in chapters 6 and 7 is only the beginning. Vector systems present a particularly interesting challenge of dimension. We must carefully capture the interactions between the system states, while accounting for the fact that stabilizing more dimensions is inherently a harder problem.

- The definitions of control capacity focus on actuation channels to provide a partial generalization for the uncertainty threshold principle. A full generalization must extend to include unreliability in the system gain A itself. The results in Chapter 4 about non-coherent estimation can be thought of as the observation counterpart to the uncertainty threshold principle, and we hope to further generalize our understanding of control capacity to include an unreliable sensing gain C.

- Calculations in Chapter 5 reveal that the scaling behavior of the control capacity of purely multiplicative noise actuation channels seems to critically depends only on the ratio of the mean to the standard deviation of the multiplicative uncertainty at "high SNR". Carry-free models provide an intuition for this, but the full story needs to be explored.

- There are specific questions to be answered about side information. We have a series of examples to show that unlike in portfolio theory, one bit of causal side information can actually increase control capacity by more than one bit! While we have a basic understanding of this in the context of carry-free models, there is more work to be done to understand this in general real-valued systems. For example, can the gain from one bit of side information in a scalar system be unbounded? What about vector systems?

- In traditional notions of information theory, the different senses of capacity are further justified by very general source-channel separation theorems. It would be interesting to see what the corresponding results are like for the information theory for robust control. Such results would require a deeper understanding of control capacity for more general systems than the linear, scalar systems considered here.

More broadly, power constraints are a type of structural limitation that is easily studied in the context of AWGN communication channels. Clearly, our frameworks must be extended to accommodate a power constraint on the control action. This can be easily done in the dynamic programming and optimization frameworks, but it will be very interesting to see what happens in the context of side information. How does this change the control capacity? Extending from these ideas, can we provide an information-theoretic view of robust control more generally?

While the work in this thesis is theoretical, the interactions between communication, control and learning are becoming increasingly relevant in practical systems. The Internet of Things (IoT) envisions an ecosystems of interacting, evolving, smart distributed controllers. Much work remains to be done to get to this vision. Research in 5G wireless networks for control are focused on providing high-reliability communication with low-latency [106]. A full understanding of the differences between control channels and communication channels can influence future system development and protocol design. For this, we first need to develop wireless communication frameworks that can capture these notions.

In addition to the obvious technical and practical questions, this dissertation opens the door for to formalize a series of questions that were harder to ask before. The work here was partly motivated by ideas in universality in both communication and compression in the information theory and computer science communities [123, 105, 118, 33]. Control provides a natural model through which we can build an information-theoretic perspective on goal-oriented communication [44]. How can we understand a notion of "universal" control, i.e. control without regard to particular system parameters or models?

In the context of portfolio theory, Cover showed that universal portfolio strategies optimization strategies exist [23]. Information theory tools can also be used to understand side information in the context of universal portfolios [24]. The natural question is: if there exists a portfolio that can perform optimally while agnostic to the parameters of the systems, under what circumstances can we design control strategies that work universally? There has also been extensive work in universal source coding (e.g. Lempel-Ziv [123]) as well as universal channel coding (e.g. the AVC approach [13, 100] or individual channels [66, 33, 76]). Can we use this information-theoretic understanding from this dissertation to develop parallel ideas for control?

Universality has also been studied in more general reinforcement learning frameworks [34] using some information-theoretic tools. The many connections between these fields need to be fleshed out so we can have a unified understanding. Reinforcement learning models consider many different types of uncertainty in systems: adversarial, truly stochastic, deterministic but unknown. Regret-based metrics have been developed to compare the performance of algorithms in the face of these uncertainties to algorithms which have access to more information. The results on the value of side information provide a parallel perspective in the linear stochastic control case. But can these notions extend and connect to understand more general systems?

Markov decision processes (MDPs) and partially observable Markov decision processes provide more general frameworks to understand decision making in systems, and it would be natural to see if we can learn to overcome structural limitations using side information ideas. Are there parallel notions of decision capacity in these settings?

This dissertation explored a range of notions of system stability. Another widely used metric for information is the log-loss metric, which looks at the log of the probability distribution of the state, and focuses on the predictability of the state. This has been recently shown to be particularly compatible with the notions of side information [56]. Using these ideas to capture the uncertainty in the system state is an interesting next step because it is a natural way to extend control capacity ideas to the discrete state setting.

Further, the work on Witsenhausen's counterexample [49, 83] immediately suggests that we should ask questions about the control capacity of decentralized systems. Information theory has a number of tools for dealing with power and information sharing across parallel channels and also has extensively developed tools in network information theory. It remains to be seen how these connect to control systems.

Finally, informational bottlenecks show up not just in communication and control, but in pretty much every single system we might consider. Economics has extensively studied the value of information and signaling in game-theoretic settings. There have historically been many far reaching connections between ideas in control, dynamic programming, optimization and game theory and economics. It would be truly elegant to have a unified understanding of actions, information, and incentives.

Bibliography

[1] Abou-Faycal, I.C. and Trott, M.D. and Shamai, S. "The capacity of discrete-time memoryless Rayleigh-fading channels". In: *Information Theory, IEEE Transactions on* 47.4 (2001), pp. 1290–1301.

[2] M. Anand and PR Kumar. "A digital interface for Gaussian relay and interference networks: Lifting codes from the discrete superposition model". In: *Arxiv preprint arXiv:1005.0167* (2010).

[3] Masanao Aoki. *Optimization of stochastic systems*. Academic Press, 1967.

[4] Himanshu Asnani and Tsachy Weissman. "On real time coding with limited lookahead". In: *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*. IEEE. 2011, pp. 60–67.

[5] Karl J Åström and Björn Wittenmark. *Adaptive control*. Courier Dover Publications, 2013.

[6] M. Athans, R. Ku, and S. Gershwin. "The uncertainty threshold principle: Some fundamental limitations of optimal decision making under dynamic uncertainty". In: *Automatic Control, IEEE Transactions on* 22.3 (1977), pp. 491–495.

[7] A.S. Avestimehr, S.N. Diggavi, and D.N.C. Tse. "Wireless network information flow: A deterministic approach". In: *Information Theory, IEEE Transactions on* 57.4 (2011), pp. 1872–1905.

[8] M Avriel and AC Williams. "The value of information and stochastic programming". In: *Operations Research* 18.5 (1970), pp. 947–954.

[9] Kerry Back and Stanley R Pliska. "The shadow price of information in continuous time decision problems". In: *Stochastics: An International Journal of Probability and Stochastic Processes* 22.2 (1987), pp. 151–186.

[10] Kristine L Bell et al. "Extended Ziv-Zakai lower bound for vector parameter estimation". In: *Information Theory, IEEE Transactions on* 43.2 (1997), pp. 624–637.

[11] Richard Ernest Bellman and Stuart E Dreyfus. "Applied dynamic programming". In: (1962).

[12] Dimitri P Bertsekas et al. *Dynamic programming and optimal control*. Vol. 1. 2. Athena Scientific Belmont, MA, 1995.

[13] D. Blackwell, L. Breiman, and AJ Thomasian. "The capacities of certain channel classes under random coding". In: *The Annals of Mathematical Statistics* 31.3 (1960), pp. 558–567.

[14] G. Bresler and D. Tse. "The two-user Gaussian interference channel: a deterministic view". In: *European Transactions on Telecommunications* 19.4 (2008), pp. 333–354.

[15] Roger W Brockett et al. "Asymptotic stability and feedback stabilization". In: (1983).

[16] Sébastien Bubeck and Nicolo Cesa-Bianchi. "Regret analysis of stochastic and non-stochastic multi-armed bandit problems". In: *arXiv preprint arXiv:1204.5721* (2012).

[17] V. Cadambe, S.A. Jafar, and S. Shamai. "Interference alignment on the deterministic channel and application to fully connected AWGN interference networks". In: *Information Theory Workshop, 2008. ITW'08. IEEE*. IEEE. 2008, pp. 41–45.

[18] Giuseppe Caire and Shlomo Shamai. "On the capacity of some channels with channel state information". In: *Information Theory, IEEE Transactions on* 45.6 (1999), pp. 2007–2019.

[19] Dan Chazan, Moshe Zakai, and Jacob Ziv. "Improved lower bounds on signal parameter estimation". In: *Information Theory, IEEE Transactions on* 21.1 (1975), pp. 90–93.

[20] Jie Chen et al. "Optimal tracking performance: preview control and exponential signals". In: *Automatic Control, IEEE Transactions on* 46.10 (2001), pp. 1647–1653.

[21] Rong-Rong Chen et al. "On fixed input distributions for noncoherent communication over high-SNR Rayleigh-fading channels". In: *Information Theory, IEEE Transactions on* 50.12 (2004), pp. 3390–3396.

[22] M. Costa. "Writing on dirty paper." In: *Information Theory, IEEE Transactions on* 29.3 (1983), pp. 439–441.

[23] Thomas M Cover. "Universal portfolios". In: *Mathematical finance* 1.1 (1991), pp. 1–29.

[24] Thomas M Cover and Erik Ordentlich. "Universal portfolios with side information". In: *Information Theory, IEEE Transactions on* 42.2 (1996), pp. 348–363.

[25] T.M. Cover and J.A. Thomas. *Elements of information theory*. Second. Wiley, 2006.

[26] M Davis. "Dynamic optimization: a grand unification". In: *Decision and Control, 1992., Proceedings of the 31st IEEE Conference on*. IEEE. 1992, pp. 2035–2036.

[27] MHA Davis. "Anticipative LQG control". In: *IMA Journal of Mathematical Control and Information* 6.3 (1989), pp. 259–265.

[28] Michael Alan Howarth Dempster. *The expected value of perfect information in the optimal evolution of stochastic systems*. Springer, 1981.

[29] Laurent El Ghaoui. "State-feedback control of systems with multiplicative noise via linear matrix inequalities". In: *Systems & Control Letters* 24.3 (1995), pp. 223–228.

[30] N. Elia. “Remote stabilization over fading channels”. In: *Systems & Control Letters* 54.3 (2005), pp. 237–249.

[31] Nicola Elia and Jeff N Eisenbeis. “Limitations of linear control over packet drop networks”. In: *Automatic Control, IEEE Transactions on* 56.4 (2011), pp. 826–841.

[32] Nicola Elia and Jeff N Eisenbeis. “Limitations of linear remote control over packet drop networks”. In: *Decision and Control, 2004. CDC. 43rd IEEE Conference on.* Vol. 5. IEEE. 2004, pp. 5152–5157.

[33] K. Eswaran et al. “Zero-rate feedback can achieve the empirical capacity”. In: *Information Theory, IEEE Transactions on* 56.1 (2010), pp. 25–39.

[34] Vivek F Farias et al. “Universal reinforcement learning”. In: *Information Theory, IEEE Transactions on* 56.5 (2010), pp. 2441–2454.

[35] SD Flåm. “Nonanticipativity in stochastic programming”. In: *Journal of Optimization theory and applications* 46.1 (1985), pp. 23–30.

[36] Emanuele Garone et al. “LQG control for distributed systems over TCP-like erasure channels”. In: *Decision and Control, 46th IEEE Conference on.* 2007, pp. 44–49.

[37] Emanuele Garone et al. “LQG control for MIMO systems over multiple erasure channels with perfect acknowledgment”. In: *Automatic Control, IEEE Transactions on* 57.2 (2012), pp. 450–456.

[38] Michael Gastpar. “To code or not to code”. PhD thesis. 2002.

[39] SI Gelfand and MS Pinsker. “Coding for a channel with random parameters”. In: *Probl. Control and Inf. Theory* 9 (1980), pp. 19–31.

[40] E Gershon, U Shaked, and I Yaesh. “Robust H-∞ estimation of stationary discrete-time linear processes with stochastic uncertainties”. In: *Systems & control letters* 45.4 (2002), pp. 257–269.

[41] Eli Gershon, Uri Shaked, and Isaac Yaesh. “¡ i¿ H¡/i¿¡ sub¿ ¡/sub¿ control and filtering of discrete-time stochastic systems with multiplicative noise”. In: *Automatica* 37.3 (2001), pp. 409–417.

[42] E Gershon et al. “Robust H-∞ filtering of stationary continuous-time linear systems with stochastic uncertainties”. In: *Automatic Control, IEEE Transactions on* 46.11 (2001), pp. 1788–1793.

[43] Richard D Gill and Boris Y Levit. “Applications of the van Trees inequality: a Bayesian Cramér-Rao bound”. In: *Bernoulli* (1995), pp. 59–79.

[44] O. Goldreich, B. Juba, and M. Sudan. “A theory of goal-oriented communication”. In: *Journal of the ACM (JACM)* 59.2 (2012), p. 8.

[45] Andrea J Goldsmith and Muriel Médard. “Capacity of time-varying channels with causal channel side information”. In: *Information Theory, IEEE Transactions on* 53.3 (2007), pp. 881–899.

[46] Andrea J Goldsmith and Pravin P Varaiya. “Capacity of fading channels with channel side information”. In: *IEEE Transactions on Information Theory* 43.6 (1997), pp. 1986–1992.

[47] P. Grover and A. Sahai. “Implicit and explicit communication in decentralized control”. In: *48th Annual Allerton Conference on Communication, Control, and Computing, Monticello, Illinois.* 2010.

[48] P. Grover and A. Sahai. “The vector Witsenhausen problem as assisted interference cancelation”. In: *International Journal on Systems, Control and Communications (IJSCC)* (2008).

[49] Pulkit Grover. “Actions can speak more clearly than words”. PhD thesis. University of California, Berkeley, 2010.

[50] Aditya Guntuboyina. “Minimax lower bounds”. PhD thesis. Yale, 2011.

[51] Vijay Gupta et al. “Receding horizon networked control”. In: *Proc. Allerton Conf. Commun., Control Comput.* 2006.

[52] V. Gupta et al. “Optimal LQG control across a packet dropping link”. In: *Automatic Control, IEEE Transactions on (Submitted).* 2004.

[53] Pushkar Hingwe et al. “Linear parameter varying controller for automated lane guidance: experimental study on tractor-trailers”. In: *Control Systems Technology, IEEE Transactions on* 10.6 (2002), pp. 793–806.

[54] Ronald A Howard. “Information value theory”. In: *Systems Science and Cybernetics, IEEE Transactions on* 2.1 (1966), pp. 22–26.

[55] Orhan C Imer, Serdar Yüksel, and Tamer Başar. “Optimal control of LTI systems over unreliable communication links”. In: *Automatica* 42.9 (2006), pp. 1429–1439.

[56] Jiantao Jiao et al. “Justification of Logarithmic Loss via the Benefit of Side Information”. In: *arXiv preprint arXiv:1403.4679* (2014).

[57] B. Juba and M. Sudan. “Universal semantic communication I”. In: *Proceedings of the 40th annual ACM symposium on Theory of computing.* ACM. 2008, pp. 123–132.

[58] Rudolph Emil Kalman et al. “A new approach to linear filtering and prediction problems”. In: *Journal of basic Engineering* 82.1 (1960), pp. 35–45.

[59] Michael Katz and Shlomo Shamai. “On the capacity-achieving distribution of the discrete-time noncoherent and partially coherent AWGN channels”. In: *Information Theory, IEEE Transactions on* 50.10 (2004), pp. 2257–2270.

[60] J Kelly Jr. “A new interpretation of information rate”. In: *Information Theory, IRE Transactions on* 2.3 (1956), pp. 185–189.

[61] Young-Han Kim. “Capacity of a Class of Deterministic Relay Channels”. In: *Information Theory, IEEE Transactions on* 54.3 (Mar. 2008), pp. 1328–1329. ISSN: 0018-9448. DOI: 10.1109/TIT.2007.915921.

[62] S.P. Kotagiri and J.N. Laneman. “Multiaccess channels with state known to some encoders and independent messages”. In: *EURASIP Journal on Wireless Communications and Networking* 2008 (2008), pp. 1–14.

[63] A. Lapidoth and S.M. Moser. “Capacity bounds via duality with applications to multiple-antenna systems on flat-fading channels”. In: *Information Theory, IEEE Transactions on* 49.10 (2003), pp. 2426–2467.

[64] A Lapidoth and S Shamai. “Fading channels: how perfect need perfect side information be?” In: *IEEE Transactions on Information Theory* 48.5 (2002), pp. 1118–1134.

[65] Xiangheng Liu and Andrea Goldsmith. “Kalman filtering with partial observation losses”. In: *Decision and Control, 2004. CDC. 43rd IEEE Conference on.* Vol. 4. IEEE. 2004, pp. 4180–4186.

[66] Y. Lomnitz and M. Feder. “Communication over individual channels”. In: *Information Theory, IEEE Transactions on* 57.11 (2011), pp. 7333–7358.

[67] Mohammad Ali Maddah-Ali and David Tse. “Completely stale transmitter channel state information is still very useful”. In: *Information Theory, IEEE Transactions on* 58.7 (2012), pp. 4418–4431.

[68] Emin Martinian and Gregory W Wornell. “Source coding with fixed lag side information”. In: *arXiv preprint cs/0410008* (2004).

[69] Emin Martinian, Gregory W Wornell, and Ram Zamir. “Source coding with distortion side information”. In: *Information Theory, IEEE Transactions on* 54.10 (2008), pp. 4638–4665.

[70] Emin Martinian, Gregory W Wornell, and Ram Zamir. “Source coding with distortion side information at the encoder”. In: *Data Compression Conference, 2004. Proceedings. DCC 2004*. IEEE. 2004, pp. 172–181.

[71] Nuno C Martins, Munther A Dahleh, and John C Doyle. “Fundamental limitations of disturbance attenuation in the presence of side information”. In: *Automatic Control, IEEE Transactions on* 52.1 (2007), pp. 56–66.

[72] Muriel Medard. “The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel”. In: *Information Theory, IEEE Transactions on* 46.3 (2000), pp. 933–946.

[73] Heinrich Meyr, Marc Moeneclaey, and Stefan A Fechtel. *Digital communication receivers: synchronization, channel estimation and signal processing, 1998.* John Wiley and Sons, New York, USA.

[74] Richard H Middleton, Jie Chen, and James S Freudenberg. “Tracking sensitivity and achievable H performance in preview control”. In: *Automatica* 40.8 (2004), pp. 1297–1306.

[75] Paolo Minero et al. “Data rate theorem for stabilization over time-varying feedback channels”. In: *Automatic Control, IEEE Transactions on* 54.2 (2009), pp. 243–255.

[76] V. Misra and T. Weissman. “The Porosity of Additive Noise Sequences”. In: *Information Theory, 2012. ISIT 2012. Proceedings. International Symposium on* (2012).

[77] Y. Mo and B. Sinopoli. “A characterization of the critical value for Kalman filtering with intermittent observations”. In: *Decision and Control, CDC. 47th IEEE Conference on*. 2008, pp. 2692–2697.

[78] A.S. Motahari, S.O. Gharan, and A.K. Khandani. “Real interference alignment with real numbers”. In: *Arxiv preprint arxiv:0908.1208* (2009).

[79] Girish N Nair and Robin J Evans. “Stabilization with data-rate-limited feedback: tightest attainable bounds”. In: *Systems & Control Letters* 41.1 (2000), pp. 49–56.

[80] Girish N Nair et al. “Feedback control under data rate constraints: An overview”. In: *Proceedings of the IEEE* 95.1 (2007), pp. 108–137.

[81] G.N. Nair and R.J. Evans. “Stabilizability of stochastic linear systems with finite feedback data rates”. In: *SIAM Journal on Control and Optimization* 43.2 (2004), pp. 413–436.

[82] U. Niesen and M.A. Maddah-Ali. “Interference alignment: From degrees-of-freedom to constant-gap capacity approximations”. In: *Information Theory (ISIT), IEEE International Symposium on*. 2012, pp. 2077–2081.

[83] Se Yong Park. “Information Flow in Linear Systems”. PhD thesis. University of California, Berkeley, Dec. 2013.

[84] Se Yong Park and Anant Sahai. “It may be easier to approximate decentralized infinite-horizon LQG problems”. In: *Decision and Control (CDC), IEEE 51st Annual Conference on*. 2012, pp. 2250–2255.

[85] S.Y. Park, G. Ranade, and A. Sahai. “Carry-free models and beyond”. In: *IEEE International Symposium on Information Theory*. 2012, pp. 1927–1931.

[86] S.Y. Park and A. Sahai. “Intermittent Kalman filtering: Eigenvalue cycles and nonuniform sampling”. In: *American Control Conference (ACC), 2011*. 2011, pp. 3692–3697.

[87] Haim H Permuter, Young-Han Kim, and Tsachy Weissman. “Interpretations of directed information in portfolio theory, data compression, and hypothesis testing”. In: *Information Theory, IEEE Transactions on* 57.6 (2011), pp. 3248–3259.

[88] Haim H Permuter, Young-Han Kim, and Tsachy Weissman. “On directed information and gambling”. In: *Information Theory, 2008. ISIT 2008. IEEE International Symposium on*. IEEE. 2008, pp. 1403–1407.

[89] Vu N Phat et al. “Robust stabilization of linear uncertain discrete-time systems via a limited capacity communication channel”. In: *Systems & control letters* 53.5 (2004), pp. 347–360.

[90] S Sandeep Pradhan, Jim Chou, and Kannan Ramchandran. "Duality between source coding and channel coding and its extension to the side information case". In: *Information Theory, IEEE Transactions on* 49.5 (2003), pp. 1181–1203.

[91] PK Rajasekaran, N Satyanarayana, and MD Srinath. "Optimum linear estimation of stochastic signals in the presence of multiplicative noise". In: *Aerospace and Electronic Systems, IEEE Transactions on* 3 (1971), pp. 462–468.

[92] Govind Ramnarayan, Gireeja Ranade, and Anant Sahai. "Side-information in Control and Estimation". In: *Information Theory, 2014. IEEE International Symposium on.* 2014.

[93] G. Ranade and A. Sahai. "Implicit communication in multiple-access settings". In: *Information Theoy, IEEE International Symposium on.* 2011, pp. 998–1002.

[94] J.S. Richters. *Communication over fading dispersive channels.* Tech. rep. DTIC Document, 1967.

[95] R Tyrrell Rockafellar and RJ-B Wets. "Nonanticipativity and L 1-martingales in stochastic optimization problems". In: *Stochastic Systems: Modeling, Identification and Optimization, II.* Springer, 1976, pp. 170–187.

[96] 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 systems". In: *Information Theory, IEEE Transactions on* 52.8 (2006), pp. 3369–3395.

[97] Anant Sahai. "Evaluating channels for control: Capacity reconsidered". In: *American Control Conference, 2000. Proceedings of the 2000.* Vol. 4. IEEE. 2000, pp. 2358–2362.

[98] Anant Sahai, Salman Avestimehr, and Paolo Minero. "Anytime communication over the Gilbert-Eliot channel with noiseless feedback". In: *Information Theory, 2005. ISIT 2005. Proceedings. International Symposium on.* IEEE. 2005, pp. 1783–1787.

[99] Anant Sahai and Sanjoy Mitter. "The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link, Part II: vector systems". In: *arXiv preprint cs/0610146* (2006).

[100] A.D. Sarwate and M. Gastpar. "Rateless codes for AVC models". In: *Information Theory, IEEE Transactions on* 56.7 (2010), pp. 3105–3114.

[101] Luca Schenato et al. "Foundations of control and estimation over lossy networks". In: *Proceedings of the IEEE* 95.1 (2007), pp. 163–187.

[102] Bruno Sinopoli et al. "Kalman filtering with intermittent observations". In: *Automatic Control, IEEE Transactions on* 49.9 (2004), pp. 1453–1464.

[103] Joel Spencer, Madhu Sudan, and Kuang Xu. "Queueing with future information". In: *ACM SIGMETRICS Performance Evaluation Review* 41.3 (2014), pp. 40–42.

[104] Ruslan Leontevich Stratonovich. "On value of information". In: *Izvestiya of USSR Academy of Sciences, Technical Cybernetics* 5 (1965), pp. 3–12.

[105] M. Sudan. “Communication amid Uncertainty”. In: ().

[106] Vasuki Narasimha Swamy et al. “Cooperative communication for high-reliability low-latency wireless control”. In: submitted to ICC 2015.

[107] Gilead Tadmor and Leonid Mirkin. “H control and estimation with preview-part II: fixed-size ARE solutions in discrete time”. In: *Automatic Control, IEEE Transactions on* 50.1 (2005), pp. 29–40.

[108] G. Taricco and M. Elia. “Capacity of fading channel with no side information”. In: *Electronics Letters* 33.16 (1997), pp. 1368–1370.

[109] S. Tatikonda. “Control under communication constraints”. PhD thesis. 2000.

[110] S. Tatikonda and S. Mitter. “Control under communication constraints”. In: *Automatic Control, IEEE Transactions on* 49.7 (2004), pp. 1056–1068.

[111] Aleksandr B Tsybakov. *Introduction to nonparametric estimation*. Springer, 2009.

[112] J Tugnait. “Stability of optimum linear estimators of stochastic signals in white multiplicative noise”. In: *Automatic Control, IEEE Transactions on* 26.3 (1981), pp. 757–761.

[113] Henry L Weidemann. *Entropy analysis of feedback control systems*. University Microfilms, 1969.

[114] Hans S Witsenhausen. “Separation of estimation and control for discrete time systems”. In: *Proceedings of the IEEE* 59.11 (1971), pp. 1557–1566.

[115] HS Witsenhausen. “A counterexample in stochastic optimum control”. In: *SIAM Journal on Control* 6 (1968), p. 131.

[116] Jacob Wolfowitz et al. *Coding theorems of information theory*. 31. Springer, 1961.

[117] Wing Shing Wong and Roger W Brockett. “Systems with finite communication bandwidth constraints I: State estimation problems”. In: *Automatic Control, IEEE Transactions on* 42.9 (1997), pp. 1294–1299.

[118] Kristen Woyach et al. “Comments on unknown channels”. In: *Information Theory Workshop (ITW), 2012 IEEE*. IEEE. 2012, pp. 172–176.

[119] Y. Wu, S. Shamai, and S. Verdú. “Degrees of freedom of the interference channel: A general formula”. In: *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*. IEEE. 2011, pp. 1362–1366.

[120] Yihong Wu and Sergio Verdú. “MMSE dimension”. In: *Information Theory, IEEE Transactions on* 57.8 (2011), pp. 4857–4879.

[121] Lihua Xie, Yeng Chai Soh, and Carlos E de Souza. “Robust Kalman filtering for uncertain discrete-time systems”. In: *Automatic Control, IEEE Transactions on* 39.6 (1994), pp. 1310–1314.

[122] Kemin Zhou and John Comstock Doyle. *Essentials of robust control*. Vol. 104. Prentice hall Upper Saddle River, NJ, 1998.

[123] Jacob Ziv and Abraham Lempel. “Compression of individual sequences via variable-rate coding”. In: *Information Theory, IEEE Transactions on* 24.5 (1978), pp. 530–536.

[124] Jacob Ziv and Moshe Zakai. “Some lower bounds on signal parameter estimation”. In: *Information Theory, IEEE Transactions on* 15.3 (1969), pp. 386–391.