Control and Optimization of Power Systems with Renewables: Voltage Regulation and Generator Dispatch

Baosen Zhang

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2013-146
http://www.eecs.berkeley.edu/Pubs/TechRpts/2013/EECS-2013-146.html

August 15, 2013

Control and Optimization of Power Systems with Renewables: Voltage Regulation and Generator Dispatch

by

Baosen Zhang

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

in

Engineering-Electrical Engineering and Computer Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor David Tse, Chair
Professor Pravin Varaiya
Professor Duncan Callaway

Fall 2013

Control and Optimization of Power Systems with Renewables: Voltage Regulation and Generator Dispatch

Abstract

Control and Optimization of Power Systems with Renewables:
Voltage Regulation and Generator Dispatch

by

Baosen Zhang

Doctor of Philosophy in Engineering-Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor David Tse, Chair

The electric power system is undergoing dramatic transformations due to the emergence of renewable resources. However, integrating these resources into the electric grid has proven to be difficult for two main reasons: these resources are uncertain and distributed. This thesis discusses how uncertainty and distributedness of the new resources are manifested as challenges in time and spatial scales, and how we can optimally control the power grid to overcome these challenges.

The theme of this thesis is that in order to control the new resources, we need a better understanding of the physical power flow in the system. To illustrate the spatial-scale challenge, we consider the voltage regulation problem for distribution networks. With a deep penetration of distributed energy resources the voltage magnitudes in a distribution system can fluctuate significantly. To control these resources such that voltage profiles remain flat, we need to coordinate the tens of thousands of households in the distribution network. By studying the geometry of power flows in the network, we show that even though the voltage regulation problem is non-convex, it can be convexified exactly through a semidefinite relaxation . Based on this insight, we an optimal and decentralized algorithm for this problem. Only communication between electrical neighbors are required in the algorithm, thus allowing it to scale to problem of large sizes.

The second problem we consider is dispatching generators in the transmission network under deep penetration of wind power. Because of on/off and ramp-rate constraints, traditional generators need to be dispatched before the actual time of delivery of energy. On the other hand, wind power cannot be dispatched and is difficult to predict in advance. This creates a mismatch in time-scale between the controls (traditional generators) and the randomness in the system (wind). We capture this challenge as a two-stage stochastic dispatch problem. In contrast to the standard Monte Carlo solutions, we show that the dimensional of the problem can be reduced dramatically by using forecast informations. This dimensional reduction is based on a better understanding of congested DC optimal power flow problems. Using this reduction, we show how to calculate the optimal reserve margin for the system in

the presence of renewables, and quantify the intrinsic impact of uncertainties on the system cost.

To my parents and Yutong Chen.

Contents

List of Figures

List of Tables

Acknowledgments

At the beginning of this dissertation, it was safe to say that I did not know much about power systems. At the end, I consider myself a "hardcore" power systems engineer. This was made possible by the many people that I had the privilege of working with, and I would like to thank them here.

First and foremost I would like to thank my advisor, David Tse, for his guidance and help throughout my graduate studies. He taught me how to do research, from choosing the right problem to writing and presenting our results. Without his enthusiasm and patience, this thesis would not have been possible. In fact, discovering and learning together with him has been the most stimulating experience of my life to date. His philosophy of looking at the simplest setting to gain intuition about the general problem will serve as the guide for my own research. I am deeply indebted to him.

I thank Prof. Pravin Varaiya and Prof. Duncan Callaway for serving on my dissertation and qual committees, and Prof. Kameshwar Poolla for serving on my qual committee.

I was very fortunate to work with many people outside of Berkeley. This thesis would be a shadow of what it is now without the collaboration with Prof. Alejandro Domínguez-García at UIUC, Prof. Ram Rajagopal at Stanford University, Prof. Javad Lavaei now at Columbia University and Prof. Albert Lam now at Baptist University of Hong Kong. Alejandro taught me much of what I know in distribution network engineering and modeling, and much of the material in Chapter 2 and 3 are the result of our joint discussions. His dedication were especially impressive since many of such discussions happened over Skype. I would like to thank Ram for introducing me to the topic of stochastic generator dispatch and Chapter 4 is the result of our joint works. His energy and openness are much appreciated, and I have learned a tremendous amount about dynamic programming and stochastic control from him. My work with Javad was done at a time when he was a postdoc at Stanford and David and I was visiting. I am deeply impressed with his technical expertise and knowledge of optimization. The results and proofs in Chapter 2 are from many stimulating discussions with him. My understanding of power flow on tree networks and OPF in general were greatly enhanced by our discussions. The work with Albert occurred when he was a postdoc in our group. I am extremely thankful for him teaching me about distributed algorithms (resulting in the algorithm in Chapter 3) and writing all the code for our simulations. I would very much like to learn his humbleness and attention to detail. I also wish to thank Prof. Andrea Montanari and Yashodhan Kanoria for their collaborations. I have learned a lot from our meetings about graph theory and probability.

I am very grateful to Eugene Litvinov's group in ISO New England for giving me the opportunity to do an internship there. I thank them for teaching me about power system market operations, and treating me as an equal member of the group. In particular, I thank Feng Zhao, Tongxin Zheng and Jinye Zhao for many helpful discussions.

My sincere thanks also goes out to my friends and colleagues. In no particular order, they are Sudeep Kamath, Changho Suh, I-Hsiang Wang, Pulkit Grover, Ka Kit Lam, Naveen Goela, Huasha Zhao, Junjie Qin, Yang Weng, Haifeng Ge, Mark Yan, Ye Sheng and Tom

Yue. I would like to especially thank Tom for his many helps, both professional and personal, over the years, dating back to our undergrad years in EngSci.

I cannot express my gratefulness to my parents in words. They have always given me their unconditional support and love. Finally, I am very lucky to have Yutong Chen in my life, who was undoubtedly my greatest discovery at Berkeley.

This thesis was partially supported by the National Science and Engineering Research Council of Canada. Canada's generosity is greatly appreciated.

Chapter 1

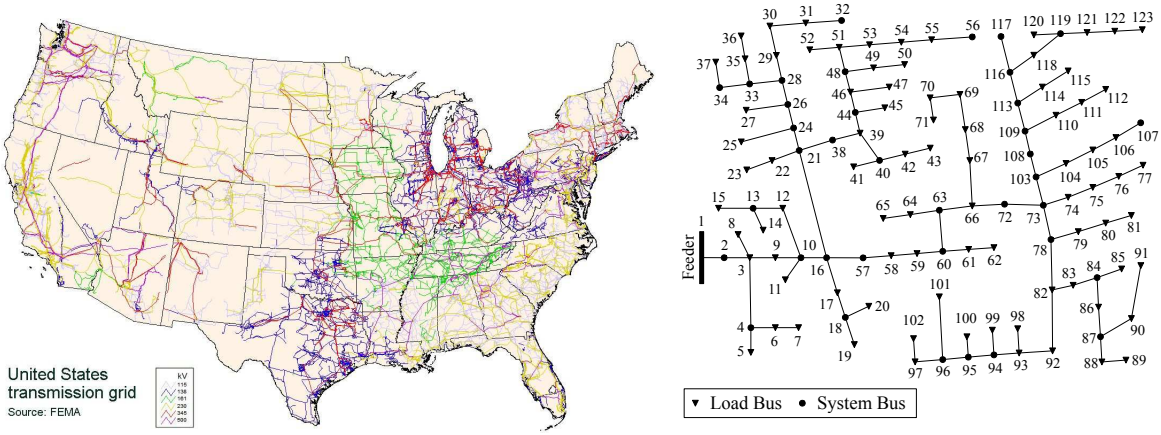
Introduction

The modern electric power system is one of the monumental engineering achievements of the twentieth century [2]. It is a vast system of interconnected machines that transfer electricity to the end consumers with a remarkable degree of reliability.[1] Yet, the power system is undergoing a dramatic transformation, both in structure and functionality. This transformation is mainly due to the emergence of a diverse range of variable generations sources, from large scale wind farms to rooftop photovoltaics (PVs). This emergence is driven by the society's desire for a more sustainable and efficient energy portfolio, often manifested as policy initiatives from various levels of government. For example, the state of California has three ambitious goals : 1) generating 33% of the state's electricity by renewables in 2030 [4]; 2) 3 GW of household solar capacity in 2017 [5]; and 3) 1.5 million electrical vehicles by 2025 [6]. Similar targets have been put in place by many other states and countries as well.

To achieve the goals set out by the various initiatives, variable renewable resources need to be integrated into the power system. This integration has proven to be non-trivial for two main reasons: the *uncertain* and *distributed* natures of these resources. These are the main difficulties since they cannot be accommodated by current operating practices, and we must design new control strategies. To describe current practices, we give a brief overview of the electricity grid. The US grid is shown in Fig. 1.1 and can be decomposed into two layers. The top layer is a meshed network that connects large generators to substations, and the bottom layer is a network that delivers electricity from the substations to the end consumers of electricity. These layers are called the transmission and distribution networks, respectively.

Currently, the main controls in the power system are the traditional generators (for example, coal, nuclear, hydro and gas plants) in the transmission network. These generators are committed and dispatched to serve the loads in the system, usually at one day ahead of actual time of operation because of on/off and ramping constraints of the generators.

[1] The planning criteria is a customer should experience no more than one hour of outage every ten years. In practice, the average US consumer experiences about 2.5 hours of blackout a year, and the average outage duration has been increasing recently because of an aging infrastructure [3].



(a) The US Transmission Network (b) Example of a Distribution Network

Figure 1.1: The transmission and distribution networks. Transmission networks have a meshed topology and operate above 100 KV; distribution networks have a tree topology and operate at much lower voltages.

The loads are the substations, that is, they are aggregations of households and industrial customers. The uncertainty nature of renewable resources means that they cannot be dispatched, that is, a wind farm cannot guarantee its output ahead of time since it is inherently stochastic. On the other hand, rest of the of the traditional (dispatchable) generators have limited ramp rates and cannot adaptively change their output to accommodate the changes in renewable output levels. Therefore, here is a *mismatch between the timescale* of control and the time scale of the randomness in the system.

The distribution network is largely ignored in real-time operations today because there is essentially nothing to control and optimize in the current distribution network. However, with the emergence of rooftop PVs and electric vehicles (EVs), each customer could be a point of control. In contrast to traditional generators, the new resources in the distribution network are controlled by their power electronics, which can respond in a time-scale of microseconds. Therefore the resources can adapt to the changes in the grid (e.g. solar radiation variation, which occurs on a minute-by-minute scale). The challenge in control is in the *spatial-scale*, since the resources are distributed over the entire network, which could contain tens of thousands of households.

The topic of this thesis is to optimally control the resources in the electricity grid under deep penetration of renewables. To illustrate the challenge in the spatial-scale, we focus on the voltage regulation problem, which is fundamental to the operation of distribution networks; to illustrate the challenges in the time-scale, we look at the stochastic dispatch of generators. Both problem can be written as optimization problems, with the former being non-convex and the latter being a high-dimensional stochastic dynamic programming problem. A central theme of this thesis is: in order to solve the optimization problem brought

on by the renewable resources, a better understanding of the fundamentals of power flow is needed. By developing fundamental insights into power flows, we show how both problem can be solved optimally and efficiently.

1.1 Problem Descriptions

1.1.1 Power Flows

Each connected power system is synchronized to a single operating frequency[2], and the voltage and current are sinusoids at this frequency; therefore, they are described by complex phasors. The interaction of voltage drops, current flows and complex power are determined by Kirchoff's first and second laws and Ohm's law [7]. Given a power system with n-buses, the physical laws can be summarized into one system of equations as:

$$\mathbf{i} = \mathbf{Y}\mathbf{v} \tag{1.1a}$$

$$\mathbf{s} = \mathbf{v} \odot \text{conj}(\mathbf{i}), \tag{1.1b}$$

where $\mathbf{v}, \mathbf{i}, \mathbf{s} \in \mathbb{C}^n$ are the voltage, current and complex power injections, respectively; $\mathbf{Y}$ is the bus admittance matrix; $\odot$ is the component-wise produce (Hadamard product) and $\text{conj}(\cdot)$ is the conjugate operator.

The real part and imaginary parts of the complex power is called active power and reactive power, labelled as $\mathbf{p} \in \mathbb{R}^n$ and $\mathbf{q} \in \mathbb{R}^n$, respectively. In practice, the possible values of active and reactive powers are restricted by the operational constraints in te system. For example, the voltage magnitude are constrained to be within a narrow range around a nominal value; the current magnitude cannot exceed thermal thresholds on the transmission lines; and complex powers are constrained at the buses. The set of all possible active power injection vectors is called the feasible injection region, denoted by $\mathcal{P}$.

Many of the problem in power system engineering can be cast as an optimization problem over $\mathcal{P}$. The most basic and the best known problem is perhaps the optimal power flow problem (OPF) in (1.2), which seeks the minimize some cost function over the set of all feasible power injections.

$$\text{minimize } \text{Cost}(\mathbf{p}) \tag{1.2a}$$

$$\text{subject to } p \in \mathcal{P}. \tag{1.2b}$$

We focus on the active powers rather than the reactive powers since cost is mainly a function of active, rather than reactive powers. The reason is that active power can do work and has unit of watt; on the other hand, reactive power cannot do work, so it has a unit of volt·ampere. For example, the electricity bill of residential customers is specified in dollars for kilowatt. This does not mean that reactive powers are not important to the operation

[2]North American grid operates at 60 Hz and European grid operates at 50 Hz. Some countries, such as Japan, have two separate grid operating at each of the frequencies, respectively.

of power systems, rather, we model reactive power as constraints that affect $\mathcal{P}$, but not the objective function. [3]

Understanding $\mathcal{P}$ is critical to solving OPF and related problems. For general networks, $\mathcal{P}$ is a complicated region in $\mathbb{R}^n$ that is not convex, and therefore (1.2) is not a convex problem. For transmission networks, $\mathcal{P}$ is often approximated to be a polytope via the DC power flow approximation (see Section 2.7 for details). This approximation is typically sufficient for modeling power flow in the transmission network, but not in the distribution network. In the latter, due to high R/X ratios and the importance of reactive power constraints, the DC approximation is not useful. For stochastic dispatch, DC power flow linearizes the problem, but the resulting stochastic program can have thousands of constraints.

1.1.2 Voltage Regulation

With the emergence of distributed energy resources (DERs) such as rooftop PV, controlling and optimization in the distribution network becomes critical to the reliable operation of the grid. One of the most important operational constraints in the distribution network is to keep the voltage magnitudes close to their reference values; and it is known that voltage magnitudes could be severely perturb by the active injections from DERs.

The voltages can be regulated by altering the active power injections of the DERs, but such changes are likely to be small. For example, a customer may agree to reduce the output of his/her rooftop solar unit by 5% to help the neighbor, but a reduction of 50% would be unlikely. Therefore instead of active power injections, the main control knobs are the *reactive power injections* provided by the power electronics on the DERs. The voltage regulation problem can be written as an OPF, where the injection region $\mathcal{P}$ is determined by: voltage, active power and reactive power constraints. A common objective function is the system loss, but other functions can be used.

The resulting OPF is not convex and is large-scale. In particular, because of the lack of communication infrastructure in the network, the problem need to be solved in a decentralized fashion.

1.1.3 Generator Dispatch

Uncertainty mainly impacts the scheduling of resources in the power system. Since most of traditional generator have on/off time constraints and limited ramping capabilities, they must be scheduled some time ahead of the actual time of electricity delivery. In North American, this is typically done one day-ahead (24 hours) of the actual time of operations. Currently, the main source of uncertainty is load, which can be predicted within 1%. Because of accurate forecasting, the scheduling of resources is solved as a deterministic OPF problem. On the other hand, wind is much more difficult to predict than load, with the state of art

[3] There are some situations where reactive power are explicitly optimized, see [8] and the references within.

forecast error being around 20% at day ahead. Under high wind penetration levels, the problem of generator scheduling can no longer be treated as deterministic.

The stochastic dispatch problem becomes a multi-period dynamic programming problem, where the traditional generator settings must be determined before the realization of the random renewable output levels are known. Curse of dimensionality becomes a major concern since typical power networks can include thousands of buses and more than ten thousand transmission lines.

1.2 Previous Works

1.2.1 Power Flow

Power flow has been studied since the 1930's. Earlier studies were done using dc network analyzer [9,10], in which the current flow was proportional to the power flow. Mathematically, this represent a linearization of the power flow equations in (1.2) and is called the DC power flow[4] (for example, see [11] and the references within). The optimal power flow problem was proposed by [12], and have received a tremendous amount of attention since then (for example, see [13–15] for surveys).

Despite the amount of research effort, OPF remains a challenge. In practice, it is generally solved using Newton-type method, which convergences to a local optimum (although convergence is not guaranteed). Some non-iterative methods have been proposed [16, 17], which may have wider radius of convergences compared to the Newton-type methods.

Relaxation methods are another class of algorithm for OPF. These algorithm converts the non-convex OPF to a convex problem through different types of relaxations. For general networks, Lagrangian dual relaxations was proposed in [18,19], conic relaxation was proposed in [20], and semidefinite programming (SDP) relaxation relaxation was proposed in [21]. The SDP relaxation technique was explored in-depth in [22] to show that the relaxation is tight for the IEEE benchmarks in [23].[5] Since the relaxations are convex, they can be solved by Newton or interior point methods in polynomial time [25]. However, the optimal solutions of relaxations are not necessarily feasible for the original OPF problem, and therefore may not be physically meaningful.

1.2.2 Voltage Regulation

In the distribution network, voltage is currently regulated through tap-changing under-load transformers, set voltage regulators, and fixed/switched capacitors. While these devices– whose operations are of mechanical nature– are effective in managing slow variations (on the time-scale of hours) in voltage, their lifetime could be dramatically reduced from the

[4]The original power flow equations are sometimes called AC power flow.

[5]Subsequent research have shown that SDP relaxation cannot be tight for all practical systems [24], and finding conditions on networks such that the relaxation is tight is an active area of research.

increased number of operations needed to handle faster voltage variations due to sudden changes (on the timescale of minutes) in active power generated or consumed by DERs.

A number of strategies have been proposed to use the power electronics on the DERs to help with voltage regulation [26–36]. These methods are either endowed with wireless or power-line communications. The general approach is to cast the voltage regulation problem as an OPF problem. A drawback of the existing methods are that they are suboptimal due to the fact that OPF problems are difficult to solve, and many require a centralized strategy (for example [33, 34]), which is difficult to implement in large networks.

1.2.3 Generator Dispatch

The study of generator scheduling is most commonly known as unit commitment and economic dispatch. Unit commitment is a integer-valued program that determines the on/off status of generators, and the subsequent OPF problem is called economic dispatch. Both problem have been studied extensively since the 1960's (for example, see [37, 38] and the references within).

Stochastic unit commitment and economic dispatch have been studied relative recently [39–44], spurred by the need to integrate wind into the system. Majority of the approaches are scenario-based, which means that the probability distribution of wind is discretized to a set of "representative" scenarios. This approaches converts the stochastic problem to a large but deterministic problems. However, the scenario based approach suffers from three disadvantages: i) what scenarios are representative? Simply finding the best discretization of the probability distribution does not always ensure the disretization is the best for the over all problem; ii) A small number of scenarios may not capture much of the stochastic information, no matter how well they are chosen; iii) Qualitative information such as the impact of uncertain on the system cost are difficult to gain using this approach. An alternative approach considering the worst case scenario using robust optimization is proposed in [45], however choosing an acceptable worst case is still difficult. In the microgrid setting, an online approach that does not require any stochastic information is proposed in [46], and a finite competitive ratio is provided.

1.3 Our Approach and Contributions

1.3.1 Power Flow

In this thesis, we take an geometric approach to understand the underlying power injection region $\mathcal{P}$. For each of the problems, we identify the critical geometric feature of the injection region that allows for them to be solved efficiently. The mathematical tools we use are convex geometry, algebraic graph theory, and linear algebra. In particular, many results in this thesis is obtained by linking the *topology* of the power network to the algebraic property of optimization problems.

In general, the power injection region $\mathcal{P}$ is known to be non-convex. We make two contribution to the understanding of $\mathcal{P}$. The first contribution is we show that even though $\mathcal{P}$ is not convex, the Pareto-front of $\mathcal{P}$ is the same as the Pareto-front of the convex hull of $\mathcal{P}$ for *power distribution networks*, under a broad set of conditions. This fact establishes that in OPF problems for distribution problems, an equivalent optimization problem can be obtained by replacing the non-convex set $\mathcal{P}$ by its convex hull, thus convexifying the OPF. Algebraically, this convex relaxation is represented by reformulating the problem as a SDP problem. The second contribution is a better understanding of the DC power flow. For transmission networks, DC power flows approximation are fairly accurate, but the resulting DC-OPF problem can still be high dimensional. Algebraically, high dimensionality corresponds the fact that $\mathcal{P}$ is described by many constraints. We show that under practical conditions, $\mathcal{P}$ is a simple geometric object, and show that it can in fact be described by a few constraints.

1.3.2 Voltage Regulation

We show that the voltage regulation problem can be thought as an OPF problem in the distribution network, and thus can be solved as an SDP. We derive a optimal decentralized algorithm that is robust to asynchrony and communication errors. The information exchange in the algorithm is between neighbors in the electric network.

1.3.3 Generator Dispatching

Using the simplified injection region for DC power flow, we show that the two-stage stochastic generator scheduling problem can be solved without resorting to scenario based techniques. Specifically, we use the first stage forecast information to reduce the dimension of second stage OPF problem. In practice, the OPF problem can be reduced to an equivalent problem for a network of a few buses (typically less than 4 for US networks), thus allowing the problem to be solved analytically. Perhaps more importantly, we characterize the *intrinsic impact of uncertainty* on the dispatch procedure, and show it is linear in the standard deviation of the uncertainty.[6]

From a higher perspective, the overall philosophy in this thesis is to first look at the simplest possible settings of problems, and once they are understood, extend them to more general situations. That is why we focus on geometry, since understanding the geometry of simple problems allows us to *guess* the correct answer to more complicated problems.

[6]Linearity is important since it allows us to answer questions of the following type: is it worth it to spend 10 million dollars to improve wind forecasts such that the uncertainty is reduced by 1%?

1.4 Thesis Outline

In Chapter 2 we introduce the topic of power flow and optimal power flow. We investigate the geometric property of the injection region by taking a step-by-step approach. We analyze the effect of each of type of constraints on the injection region, starting with the case of no constraints. We show that the injection region is the upper halfspace given the conservation of energy. Next, we focus on distribution networks. Because of its tree structure, we show the injection region can be thought as a positive linear transformation of the product of two-bus injection regions. Therefore all convexity results about two-bus networks would carry over to the general tree network. We show that injection regions for two-bus networks are a partial ellipse, where the Pareto-front is preserved under the convex hull operation. We then give an numerical algorithm for the convex relaxation. We finish this chapter by describing the DC flow approximation.

In Chapter 3 we solve the voltage regulation problem by casting it as an OPF. Moreover, we provide an optimal decentralized algorithm that only require communication between the neighboring buses. The information exchanged are a set of Lagrange multipliers, and each bus does not need to reveal its private data. This algorithm is shown to be tolerant to communication errors and asynchrony through numerical simulations.

In Chapter 4 we introduce solve the two-stage generator dispatch problem. The main idea is to use the first stage prediction to simplify the second stage DC-OPF. Namely, we show that even though the number of transmission constraints are large, only a few are binding at a time. Using forecast information, we can predict the binding constraints and the transmission capacity limits on other lines can be ignored. This allows us to reduce the dimension of the OPF problem dramatically and solve the stochastic dispatch problem in an analytical fashion.

Chapter 5 concludes this thesis, and the appendices contains most of the proofs.

Chapter 2

Optimal Power Flow

In this chapter, we first provide the power flow equations that describes the power systems that we consider in this thesis, then define the optimal power flow problem. We then consider three different scenarios where the optimal power flow can be solved *exactly*, they are i) a system without voltage constraints; ii) networks that have a tree topology (distribution networks); and iii) lossless networks with that is combination of cycles and trees. The DC power flow is also introduced at the end of the chapter.

2.1 Power Flow Equations

Consider an electric network with n buses. Throughout we assume the network is connected. We write $i \sim k$ if bus i is connected to k, and $i \nsim k$ if they are not connected. Let z_{ik} denote the complex impedance of the transmission line between bus i and bus k, and $y_{ik} = \frac{1}{z_{ik}} = g_{ik} + jb_{ik}$. We have $g_{ik} > 0$, and we assume that the lines are inductive (as in the Pi model) so $b_{ik} < 0$. Note that $z_{ik} = z_{ki}$ and $y_{ik} = y_{ki}$. Let z_{ii} (y_{ii}) denote the shunt impedance (admittance) of bus i to ground. These shunt impedances can come from the capacitance to ground in the Pi model of the transmission line, the capacitor banks installed for reactive power injection, or modeling constant impedance loads. The bus admittance matrix is denoted by $\mathbf{Y}$ and defined as

$$Y_{ik} = \begin{cases} \sum_{l \sim i} y_{il} + y_{ii} & \text{if } i = k \\ -y_{ik} & \text{if } i \sim k \\ 0 & \text{if } i \nsim k \end{cases}. \tag{2.1}$$

$\mathbf{Y}$ is symmetric. If the entries of $\mathbf{Y}$ are real, we say the network is purely resistive and if the entries are imaginary, we say the network is lossless. Lines in the transmission network are mainly inductive so it is sometimes assumed that the network is lossless.

Let $\mathbf{v} = [V_1\ V_2\ \dots\ V_n]^T \in \mathbb{C}^n$ be the vector of bus voltages and $\mathbf{i} = [I_1\ I_2\ \dots\ I_n]^T \in \mathbb{C}^n$

be the vector of currents[1], where I_i is the total current flowing out of bus i to the rest of the network. By Ohm's law and Kirchoff's Current Law,

$$\mathbf{i} = \mathbf{Y}\mathbf{v} \tag{2.2a}$$

$$\mathbf{s} = \mathbf{v} \odot \text{conj}(\mathbf{i}), \tag{2.2b}$$

where $\mathbf{s} = [S_1 \ S_2 \dots S_n]^T \in \mathbb{C}^n$ is the complex power;$\odot$ is the component-wise produce (Hadamard product) and conj$(\cdot)$ is the conjugate operator. In rectangular coordinates, $S_i = P_i + jQ_i$ where P_i is the active power and Q_i is the reactive power. A positive P_i means bus i is generating active power and a negative P_i means bus i is consuming active power; similarly for reactive powers. Let $\mathbf{p} = (P_1 \ P_2 \ \dots \ P_n)$ be the vector of active powers and $\mathbf{q} = (Q_1 \ Q_2 \ \dots \ Q_n)$ be the vector of reactive powers.

Eliminating $\mathbf{i}$ from (2.2), we obtain

$$\mathbf{s} = \mathbf{p} + j\mathbf{q} = \text{diag}(\mathbf{v}\mathbf{v}^H\mathbf{Y}^H), \tag{2.3}$$

where diag$(\cdot)$ is the diagonal operator[2] and $(\cdot)^H$ is the Hermitian transpose. The resistive loss on the transmission line between buses i and k is given by

$$L_{ik} = |V_i - V_k|^2 g_{ik};$$

the active and reactive powers flowing from bus i to k are denoted as P_{ik} and Q_{ik}, given by

$$P_{ik} + jQ_{ik} = V_i|V_i - V_k|^H y_{ik}^H.$$

Note $L_{ik} = P_{ik} + P_{ki}$.

2.2 Optimal Power Flow

A number of optimization problems in power systems can be written as the following OPF problem:

$$\text{minimize } f(P_1, P_2, \dots, P_n) \tag{2.4a}$$

$$\text{subject to } \underline{V}_i \le |V_i| \le \overline{V}_i \tag{2.4b}$$

$$L_{ik} \le l_{ik} \tag{2.4c}$$

$$P_{ik} \le \overline{P}_{ik} \tag{2.4d}$$

$$\underline{P}_i \le P_i \le \overline{P}_i \tag{2.4e}$$

$$\underline{Q}_i \le Q_i \le \overline{Q}_i \tag{2.4f}$$

$$\mathbf{p} + j\mathbf{q} = \text{diag}(\mathbf{v}\mathbf{v}^H\mathbf{Y}^H), \tag{2.4g}$$

[1]Following the convention in power engineering, scalars representing voltage, current and power are denoted with capital letters. We use $\mathbf{x}$ to denote vectors, and $\mathbf{X}$ to denote matrices.

[2]This operator is overloaded as it is in Matlab. Given a vector, diag returns a diagonal matrix; and given square matrix, diag returns diagonal of the matrix as a column vector

where $f(P_1, P_2, \ldots, P_n)$ is the cost function (not necessarily quadratic) defined on the real powers; (2.4b), (2.4c), (2.4d), (2.4e) and (2.4f) are the constraints corresponding to bus voltage, line thermal loss, line power flow and bus real and reactive power respectively; and (2.4g) is the physical law coupling voltage to power. The thermal loss constraints in (2.4c) are calculated from current rating of transmission lines and are usually the dominant constraints in distribution networks [47]. Typically the data sheet of a line would have a maximum current rating $I_{\max}$ of the line, and this gives $l_{ik} = I_{\max}^2 R$, the maximum loss that can be tolerated across a line. In practice, f is usually an increasing function of the power injections. For example, if $f(P_1, \ldots, P_n) = P_1 + \cdots + P_n$, then we are minimizing the loss in the network; or if f is quadratic with positive coefficients, then we are minimizing the cost of generation.

A central object of study in this thesis is the feasible injection region, $\mathcal{P}$, defined as

$$\begin{aligned}\mathcal{P} = \{\mathbf{p} \in \mathbb{R}^n : \mathbf{p} = \mathrm{Re}(\mathrm{diag}(\mathbf{v}\mathbf{v}^H\mathbf{Y}^H)), \underline{V}_i \leq |V|_i \leq \overline{V}_i \ \forall i,\\ L_{ik} \leq l_{ik} \ \forall i \sim k, P_{ik} \leq \overline{P}_{ik} \ \forall i \sim k, \underline{P}_i \leq P_i \leq \overline{P}_i \forall i, \underline{Q}_i \leq Q_i \leq \overline{Q}_i\}. \end{aligned} \tag{2.5}$$

Therefore $\mathcal{P}$ is the feasibility region of (2.4), and we can rewrite (2.4) as

$$\text{minimize } \mathbf{f}(\mathbf{p}) \tag{2.6a}$$
$$\text{subject to } p \in \mathcal{P}. \tag{2.6b}$$

Note the reactive powers are represented as a constraint of the injection region. This is because in most practical settings, the objective function of the optimization problem is in terms of *real* powers only. For example, the cost curve for an generator only includes the real power output; also, the consumers are only charged based on the amount of real power they consume (watt-hours). Since the objective function is in terms of real powers only, the injection region is the set of all real injections.

Since the objective function in (2.4) and (2.6) is generally convex, the convexity of the OPF depend on the geometric property of the region $\mathcal{P}$. In rest of the section, we establish conditions on the power systems such that the OPF problem is convex. An important note is that $\mathcal{P}$ *need not necessarily be convex*. Furthermore, we provide efficient algebraic descriptions of $\mathcal{P}$.

2.3 Network with No Operation Constraints

To warm up, let us first consider a network with no operation constraints. Since there are no constraints, the injection region is defined as

$$\mathcal{P} = \{\mathbf{p} \in \mathbb{R}^n : \mathbf{p} = \mathrm{Re}(\mathrm{diag}(\mathbf{v}\mathbf{v}^H\mathbf{Y}^H)). \tag{2.7}$$

The reactive powers are ignored since we model reactive power as constraints in (2.4). In this case, the injection region has a simple characterization.

Theorem 1. *If the network is lossy*[3]*, then $\mathcal{P}$ is given by*

$$\mathcal{P} = \{\mathbf{p} \in \mathbb{R}^n : \sum_{i=1}^{n} P_i > 0\} \cup \{\mathbf{0}\}. \tag{2.8}$$

Therefore $\mathcal{P}$ is the union of the open upper half space of $\mathbb{R}^n$ and the origin $\mathbf{0}$. Note this region is connected and convex. If the network is lossless, then $\mathcal{P}$ is given by

$$\mathcal{P} = \{\mathbf{p} \in \mathbb{R}^n : \sum_{i=1}^{n} P_i = 0\}. \tag{2.9}$$

Therefore $\mathcal{P}$ is a hyperplane through the origin.

This result is intuitive pleasing since it says if there are no constraints in the network then the injection region is only limited by the law of conservation of energy. Conservation of energy gives the bound $\sum_{i=1}^{n} P_i \geq 0$, and if the network is not lossless then $\sum_{i=1}^{n} P_i > 0$ except when all voltages are equal. In this case, all injections are 0 so $\mathbf{p} = \mathbf{0}$. Theorem 1 states this is the only constraint on the injection region. The authors in [48, 49] conjectured that the unconstrained injection region is convex, and (2.8) shows this is indeed the case. To proof this theorem, it is necessary to show that for every vector $\mathbf{p} \in \mathcal{P}$, there exists a voltage $\mathbf{v}$ that achieves $\mathbf{p}$. The details are given in the Appendix A.

In practice, some of the constraints in (2.4) would be binding. For example, the voltages magnitudes at each bus are bounded. Figure 2.3 shows the injection region of a two bus network with fixed voltage magnitudes. The region is an ellipse (without the interior). The next few sections is devoted to the study of the effect of constraints on the injection regions of tree networks and their implications to optimization problems.

2.4 Distribution Networks

Power systems are operated as radial networks [47], therefore we model it as a connected tree with vertex set $\mathcal{V} = \{1, \ldots, n\}$ and the edge set $\mathcal{E}$.[4] The notation $(i, k) \in \mathcal{E}$ implies that there exists a line connecting bus i and bus k. Figure 2.1 is an example of a commonly used distribution test system, the 123-bus network [50]. This section investigates the properties of the power injection region $\mathcal{P}$ for tree networks.

2.4.1 Two-Bus Networks

Consider the two-bus network in Figure 2.2 with the line admittance $g - jb$. Let the com-

[3]Every line has non-zero resistance

[4]A distribution system may contain cycles. However, at anytime of operation, the switches in the system are opened or closed such that the network is a tree. This design is used because of its simplicity in fault detection and isolation.

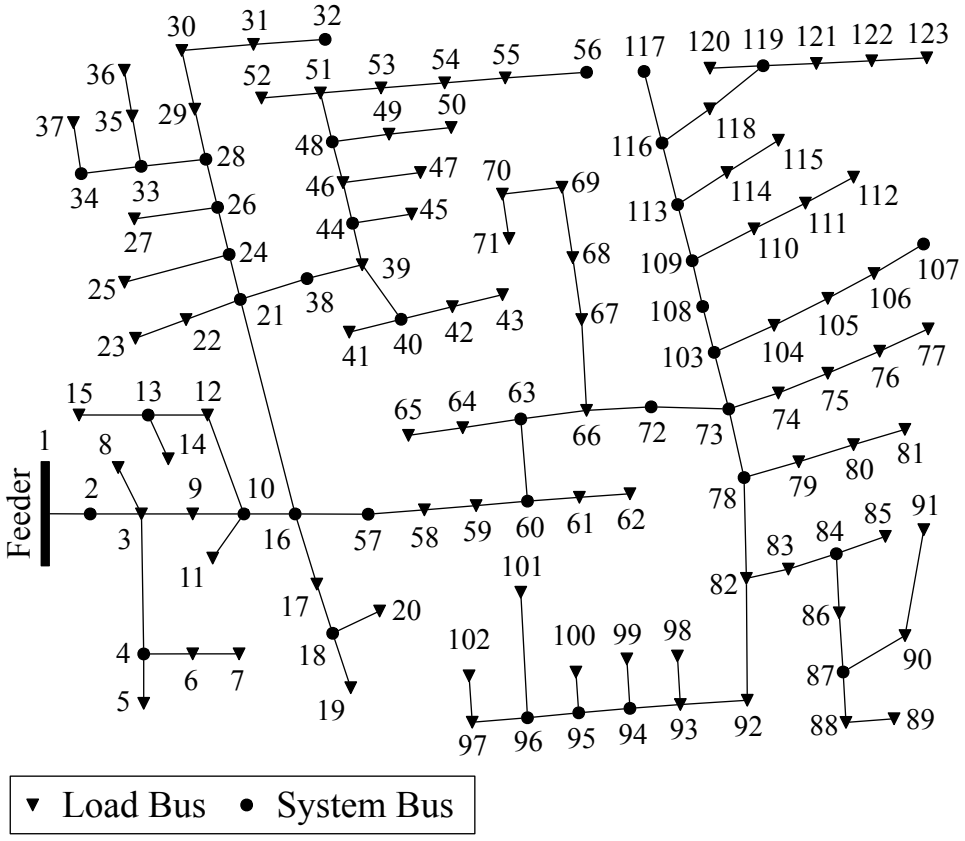


Figure 2.1: 123-bus system: electrical network graph.

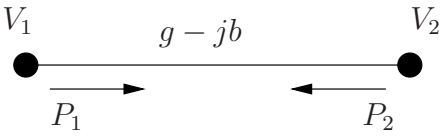


Figure 2.2: A two-bus network.

plex voltages at buses 1 and 2 be expressed as $V_1 = |V_1| \exp(j\theta_1)$ and $V_2 = |V_2| \exp(j\theta_2)$. Throughout this subsection, assume that the magnitudes $|V_1|$ and $|V_2|$ are fixed, while θ_1 and θ_2 are variable. The power injections at the two buses are given by

$$P_1 = |V_1|^2 g + |V_1||V_2| b \sin(\theta) - |V_1||V_2| g \cos(\theta) \tag{2.10a}$$

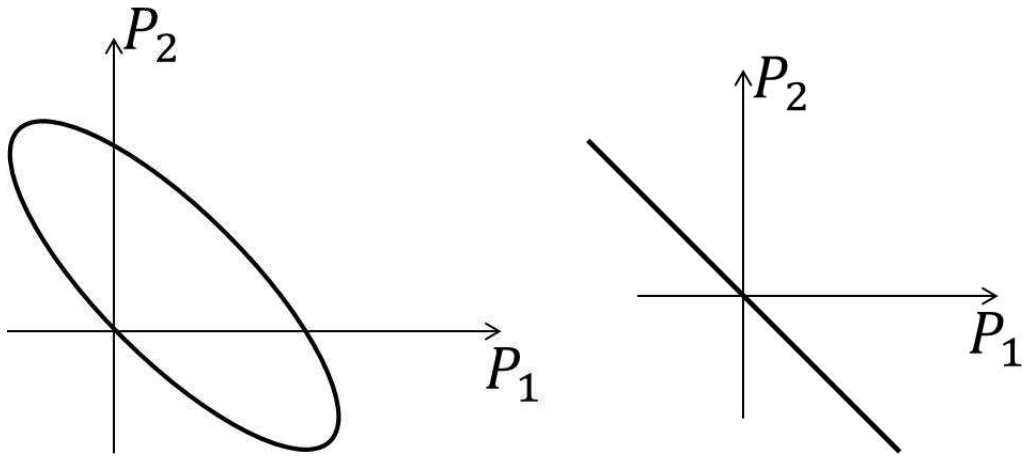
$$P_2 = |V_2|^2 g - |V_1||V_2| b \sin(\theta) - |V_1||V_2| g \cos(\theta), \tag{2.10b}$$

where $\theta = \theta_1 - \theta_2$. Since the network has only two buses, $P_1 = P_{12}$ and $P_2 = P_{21}$, where P_{ik} is the power flowing out of bus i to bus k. Since the voltage magnitudes are fixed, the power flows between the buses can both be described in terms of the single parameter θ. Notice

that a circle centered at the origin and of radius 1 can be parameterized as $(\cos(\theta), \sin(\theta))$. Therefore, (2.10) represents an affine transformation of a circle, which leads to an ellipse. This ellipse contains all points (P_1, P_2) satisfying the inequality

$$\left\| \begin{bmatrix} b & -g \\ -b & -g \end{bmatrix}^{-1} \begin{bmatrix} P_1 - |V_1|^2 g \\ P_2 - |V_2|^2 g \end{bmatrix} \right\|_2 = |V_1||V_2|,$$

where $\|\cdot\|_2$ denotes the 2-norm operator. As can be seen from the above relation, the ellipse is centered at $(|V_1|^2 g, |V_2|^2 g)$, where its major axis is at an angle of -45° to the x-axis with length $|V_1 V_2| b$ and its minor principle axis has length $|V_1 V_2| g$. If the line is lossy, the injection region is a hollow ellipse as shown in Figure 2.3a. If the line is lossless, the ellipse is degenerate and collapses into a line through the origin as shown in Figure 2.3b. In practice



(a) lossy (b) lossless

Figure 2.3: The region defined by (2.10): (a) shows the region corresponding to $|V_1| = |V_2| = 1$ (per unit), $b = 5$ and $g = 1$; (b) shows the region for a lossless line.

most lines in distribution networks are lossy with b/g ratio typically between 3 to 5 (instead of > 10 in transmission networks) [47, 50]. Thus, the interesting and practical case is when the region is a hollow ellipse. Note that the convex hull of this region is the filled ellipse.

Now, we investigate the effect of thermal, line flow and angle constraints. Since the network has fixed voltage magnitudes, the thermal loss and line flow constraints can be recast as angle constraints of the form $\underline{\theta} \le \theta \le \overline{\theta}$ for some limits $\underline{\theta} \in [-\pi, 0]$ and $\overline{\theta} \in [0, \pi]$. More precisely, the loss of the line, denoted by L_{12}, can be calculated as

$$\begin{aligned} L_{12} &= |V_1 - V_2|^2 g = P_{12} + P_{21} \\ &= |V_1|^2 g - 2|V_1||V_2| g \sin(\theta) + |V_2|^2 g. \end{aligned} \tag{2.11}$$

It follows from the above equality that a loss constraint $L_{12} \le \overline{L}_{12}$ (for a given $\overline{L}_{12}$) can be translated into an angle constraint. Likewise, the line flow inequalities $P_{12} \le \overline{P}_{12}$ and

$P_{21} \leq \overline{P}_{21}$ are also angle constraints. As a result, we restrict our attention only to angle constraints in the rest of this part.

We define the injection region $\mathcal{P}$ to be the set of all points $\{(P_1, P_2)\}$ given by (2.10) by varying $\theta \in [\underline{\theta}, \overline{\theta}]$. The bold curve in Figure 2.4 represents the injection region after a certain angle constraint.

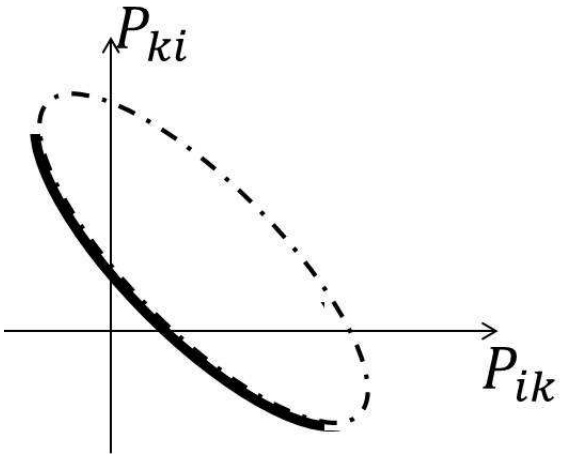


Figure 2.4: The feasible set for the two flows along a line when there are power flow constraints. It is a subset of an ellipse which is the feasible set when there are no constraints other than the fixed voltage magnitudes at the two buses. In this example, the feasible set is part of the Pareto-front of the ellipse.

The key object in this section is the *Pareto-front of a set*. Given a set $\mathcal{A} \subset \mathbb{R}^n$, let $\text{conv}(\mathcal{A})$ denote the convex hull of $\mathcal{A}$. A point $\mathbf{x} \in \mathcal{A}$ is Pareto-optimal if there does not exist another point $\mathbf{y} \in \mathcal{A}$ such that $\mathbf{y} \leq \mathbf{x}$ with strict inequality in at least one coordinate.[5] Let $\mathcal{O}(\mathcal{A})$ denote the set of all Pareto-optimal points of $\mathcal{A}$, which is called the Pareto-front of $\mathcal{A}$. Note that if a strictly increasing function is minimized over $\mathcal{A}$, its optimal solution must belong to $\mathcal{O}(\mathcal{A})$ [51]. The important property of the non-convex feasible set $\mathcal{P}$ for a two-bus network is that Pareto front of P is the same as the Pareto-front of the convex hull of $\mathcal{P}$ (see Figure 2.4). To understand the usefulness of this property in solving an optimization problem over this region, consider the following pair of optimization problems for a strictly increasing function f:

$$\text{minimize } f(P_1, P_2) \tag{2.12a}$$
$$\text{subject to } (P_1, P_2) \in \mathcal{P}, \tag{2.12b}$$

and

$$\text{minimize } f(P_1, P_2) \tag{2.13a}$$
$$\text{subject to } (P_1, P_2) \in \text{conv}(\mathcal{P}). \tag{2.13b}$$

[5]Given to vectors, $\mathbf{x} \leq \mathbf{y}$ means that $x_i \leq y_i$ for all i.

Since f is strictly increasing in both of its arguments, the optimal solution to (2.13) must be on the Pareto boundary of the feasible set; therefore both optimization problems share the same solution $(P_1^*, P_2^*) \in \mathcal{P}$. This implies that instead of solving the non-covex problem (2.12), one can equivalently solve the optimization (2.13) that is always convex for a convex function f. Hence, even though $\mathcal{P}$ is not convex, optimization over $\mathcal{P}$ and $\text{conv}(\mathcal{P})$ is equivalent for a broad range of optimization problems due to the following lemma.

Lemma 2. *Let $\mathcal{P} \in \mathbb{R}^2$ be the two-bus injection region defined in* (2.10) *by varying θ over $[\underline{\theta}, \overline{\theta}]$. The relation $\mathcal{O}(\mathcal{P}) = \mathcal{O}(\text{conv}(\mathcal{P}))$ holds.*

2.4.2 Tree Networks with Line Constraints

In this subsection, we extend Lemma 2 to an arbitrary tree network with constraints on the transmission lines (for example, thermal and line flow constraints).

First, we express the injection region of a general tree as a *linear transformation* of the power flow region. Given a general network described by its admittance matrix $\mathbf{Y}$, consider a connected pair of buses i and k. Let P_{ik} denote the power flowing from bus i to bus k through the line (i, k) and P_{ki} denote the power flowing from bus k to bus i. Similar to the two-bus case studied earlier, one can write:

$$\begin{aligned} P_{ik} &= |V_i|^2 g_{ik} + |V_i||V_k| b_{ik} \sin\theta_{ik} - |V_i||V_k| g_{ik} \cos\theta_{ik} \\ P_{ki} &= |V_k|^2 g_{ik} - |V_i||V_k| b_{ik} \sin\theta_{ik} - |V_i||V_k| g_{ik} \cos\theta_{ik}, \end{aligned}$$

where $\theta_{ik} = \theta_i - \theta_k$. The tuple (P_{ik}, P_{ki}) is referred to as the flow on the line (i, k). As in the two-bus case, all the thermal and line flow constraints can be cast as a constraint on the angle θ_{ik}. Note that the angle constraint on θ_{ik} only affects the flow on the line (i, k); therefore it is called a local constraint.

There are $2|\mathcal{E}|$ numbers describing the flows in the network. Let $\mathcal{F}$ denote the feasible set of the flows in $\mathbb{R}^{2|\mathcal{E}|}$, where the bus voltage magnitudes are fixed across the network and each flow satisfies its local constraints. Recall that the net injection at bus i is related to the line flows through the relation $P_i = \sum_{k:k\sim i} P_{ik}$. This motivates the introduction of an $n \times 2|\mathcal{E}|$ matrix $\mathbf{A}$ defined below with rows indexed by the buses and the columns indexed by the lines:

$$A(i, (k, l)) = \begin{cases} 1 & \text{if } i = k \\ 0 & \text{otherwise.} \end{cases} \tag{2.15}$$

The matrix $\mathbf{A}$ can be seen as a generalization of the edge-to-node adjacency matrix of the graph. The injection vector $\mathbf{p} \in \mathbb{R}^n$ and the flow vector $\mathbf{f} \in \mathcal{F}$ are related by $\mathbf{p} = \mathbf{A}\mathbf{f}$. We express the set of line flows as $\{P_{ik}, P_{ki}\}$ and say that $\mathbf{p}$ is achieved by the set of flows. This implies that the feasible injection region $\mathcal{P}$ is given by

$$\mathcal{P} = \mathbf{A}\mathcal{F}. \tag{2.16}$$

Since the above mapping is linear, it is straightforward to show that $\text{conv}(\mathcal{P}) = \mathbf{A}\,\text{conv}(\mathcal{F})$.

We now demonstrate that $\mathcal{F}$ has a very simple structure: it is simply a *product* of the two-bus flow regions, one for each line in the network:

$$\mathcal{F} = \prod_{(i,k)\in\mathcal{E}} \mathcal{F}_{ik}, \tag{2.17}$$

where the two-dimensional set $\mathcal{F}_{ik}$ is the two-bus flow region of the line (i,k). In other words, the flows along different lines are decoupled. To substantiate this fact, it suffices to show that the flow on an arbitrary line of the network can be adjusted without affecting the flows on other lines. To this end, consider the line (i,k) and a set of voltages with the angles $\theta_1,\ldots,\theta_n$. The power flow along the line (i,k) is a function of $\theta_{ik} = \theta_i - \theta_k$. Assume that we want to achieve a new flow on the line associated with some angle $\tilde{\theta}_{ik}$. In light of the tree structure of the network, it is possible to find a new set of angles $\tilde{\theta}_1,\ldots,\tilde{\theta}_n$ such that $\tilde{\theta}_i - \tilde{\theta}_k = \tilde{\theta}_{ik}$ and that the angle difference is preserved for every line in $\mathcal{E}\backslash(i,k)$.

Due to this product structure of $\mathcal{F}$, it is possible to generalize Lemma 2.

Lemma 3. *Given a tree network with fixed voltage magnitudes and local angle constraints, consider the injection set $\mathcal{P}$ defined in* (2.16)*. The relation $\mathcal{O}(\mathcal{P}) = \mathcal{O}(\mathrm{conv}(\mathcal{P}))$ holds.*

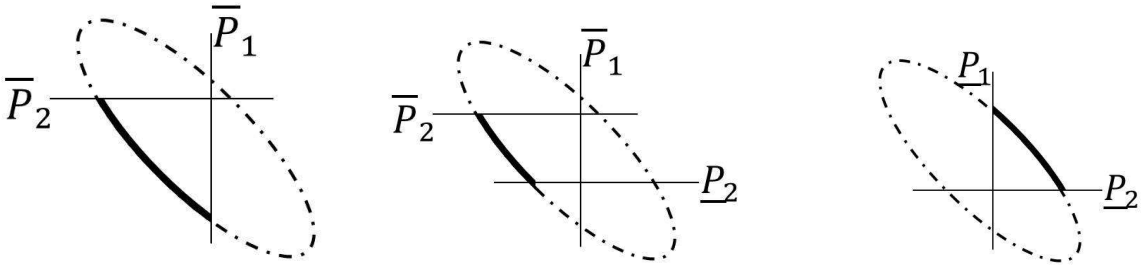
Proof: First, we show that $\mathcal{O}(\mathrm{conv}(\mathcal{P})) \subseteq \mathcal{O}(\mathcal{P})$. Given $\mathbf{p} \in \mathcal{O}(\mathrm{conv}(\mathcal{P}))$, let $\{(P_{ik}, P_{ki})\} \in \mathrm{conv}(\mathcal{F})$ be the set of flows that achieves $\mathbf{p}$. Consider a line $(i,k) \in \mathcal{E}$. Since $\mathcal{F}$ is a product space and $\mathbf{p} \in \mathcal{O}(\mathrm{conv}(\mathcal{P}))$, we have $(P_{ik}, P_{ki}) \in \mathcal{O}(\mathrm{conv}(\mathcal{F}_{ik}))$. Moreover, it follows from Lemma 2 that $\mathcal{O}(\mathrm{conv}(\mathcal{F}_{ik})) = \mathcal{O}(\mathcal{F}_{ik})$. Therefore, $(P_{ik}, P_{ki}) \in \mathcal{F}_{ik}$ for every line (i,k). This gives $\mathbf{p} \in \mathcal{P}$ and consequently $\mathbf{p} \in \mathcal{O}(\mathcal{P})$.

Next, we show that $\mathcal{O}(\mathcal{P}) \subseteq \mathcal{O}(\mathrm{conv}(\mathcal{P}))$. Given $\mathbf{p} \in \mathcal{O}(\mathcal{P})$, assume that $\mathbf{p} \notin \mathcal{O}(\mathrm{conv}(\mathcal{P}))$. Then, there exists a point $\mathbf{p}' \in \mathcal{O}(\mathrm{conv}(\mathcal{P}))$ such that $\mathbf{p}' \leq \mathbf{p}$ with strict inequality in at least one coordinate. By the first part of the proof, we have $\mathbf{p}' \in \mathcal{P}$, which contradicts $\mathbf{p} \in \mathcal{O}(\mathcal{P})$. ■

2.4.3 Two-Bus Network with Bus Power Constraints

So far, we have studied tree networks with angle constraints and global bus power constraints. Power constraints are harder to deal with since they represent *global constraints*: they couple all the flow into and out of a bus. In contrast, angle constraints are local in the sense they are independent across the transmission lines. We want to investigate the effect of bus power constraints. We first consider the two-bus network shown in Figure 2.2, and incorporate the angle constraints together with the bus active power constraints of the form $\underline{P}_i \leq P_i \leq \overline{P}_i$ for $i = 1, 2$. Let $\mathcal{P}_\theta = \{(P_1,P_2) : |V_1| = \overline{V}_1,\ |V_2| = \overline{V}_2,\ \underline{\theta} \leq \theta \leq \overline{\theta}\}$ be the angle-constrained injection region, and $\mathcal{P}_P = \{(P_1,P_2) : \underline{P}_i \leq P_i \leq \overline{P}_i,\ i = 1,2\}$ be the bus power constrained region, where $\overline{V}_1$ and $\overline{V}_2$ are the given nominal values of the voltage magnitudes. The overall injection region is given by the intersection of the two regions through the equation

$$\mathcal{P} = \mathcal{P}_\theta \cap \mathcal{P}_P. \tag{2.18}$$



(a) (b) (c)

Figure 2.5: Three possible cases for the bus power constrained injection region.

There are several possibilities for the shape of $\mathcal{P}$, as visualized in Figures 2.5a, 2.5b and 2.5c. In Figure 2.5a, both buses have power upper bounds. In Figure 2.5b, P_1 has upper bound, while P_2 has both upper and lower bounds. In Figure 2.5c, both buses have lower bounds. It can be observed that $\mathcal{O}(\mathcal{P}) = \mathcal{O}(\text{conv}(\mathcal{P}))$ for Figures 2.5a-2.5b, but this desirable property does not hold for Figure 2.5c.

Figure 2.5c means that in the presence of active power lower bounds, the relationship $\mathcal{O}(\mathcal{P}) = \mathcal{O}(\text{conv}(\mathcal{P}))$ does not always hold. This is the reason for the various assumptions made about bus power lower bounds in [52–54]. Note that θ_{12} in Figure 2.5c is allowed to vary from $-\pi$ to π. However, the angles are often constrained in practice by thermal and/or stability conditions. For example, the thermal constraints usually limit the angle difference on a line to be less than 10°. Figure 2.6 shows a typical distribution network together with its thermal constraints, from which it can be observed that each angle difference is restricted to be less than 7°. Flow constraints also limit the angle differences in a similar fashion.

Assume that the angle constraints are such that $\mathcal{P}_\theta = \mathcal{O}(\text{conv}(\mathcal{P}_\theta))$, implying that every point in $\mathcal{P}_\theta$ is Pareto-optimal. Now, there are two possible scenarios for the injection region $\mathcal{P}$ as shown in Figure 2.7. In Figure 2.7a, some of the points of the region $\mathcal{P}_\theta$ remain in $\mathcal{P}$ and they form the Pareto-front of both $\mathcal{P}$ and conv($\mathcal{P}$). In Figure 2.7b, $\mathcal{P} = \emptyset$ so then conv($\mathcal{P}$) = $\emptyset$ as well. We observe in both cases that $\mathcal{O}(\mathcal{P}) = \mathcal{O}(\text{conv}(\mathcal{P}))$. Therefore, we have $\mathcal{O}(\mathcal{P}) = \mathcal{O}(\text{conv}(\mathcal{P}))$ if $\mathcal{P}_\theta = \mathcal{O}(\text{conv}(\mathcal{P}_\theta))$.

In terms of the line parameters b_{12} and g_{12}, the condition $\mathcal{P}_\theta = \mathcal{O}(\text{conv}(\mathcal{P}_\theta))$ can be written as:

$$-\tan^{-1}(\frac{b_{12}}{g_{12}}) < \underline{\theta}_{12} \leq \overline{\theta}_{12} < \tan^{-1}(\frac{b_{12}}{g_{12}}). \tag{2.19}$$

Observe that $\tan^{-1}(b_{12}/g_{12})$ is equal to 45.0°, 63.4° and 78.6° for $\frac{b_{12}}{g_{12}}$ equal to 1, 2 and 5, respectively. These numbers suggest that the above condition is very practical. Note that since the inductance of a practical AC transmission line is larger than its resistance, the above requirement is met if $|\underline{\theta}_{12}|, |\overline{\theta}_{12}| < 45.0°$. It is noteworthy that an assumption

$$-\tan^{-1}(\frac{g_{12}}{b_{12}}) < \underline{\theta}_{12} \leq \overline{\theta}_{12} < \tan^{-1}(\frac{g_{12}}{b_{12}}) \tag{2.20}$$

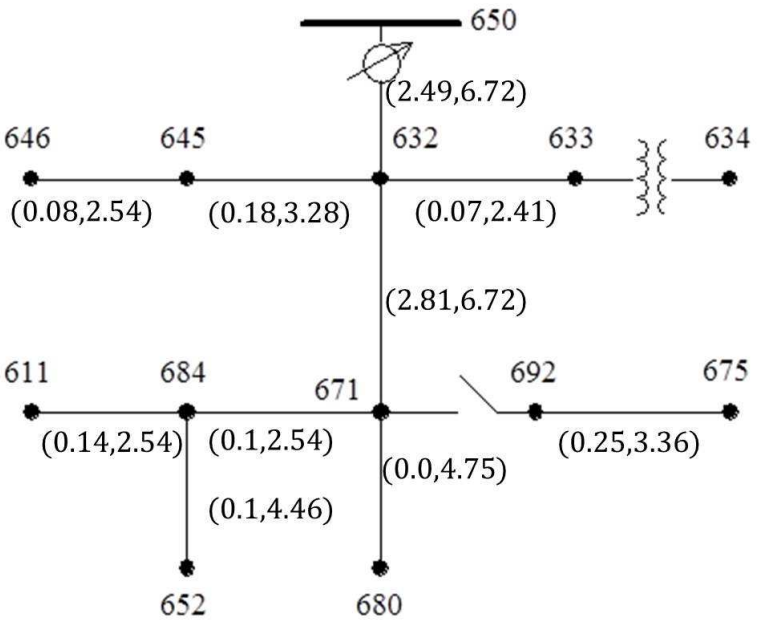
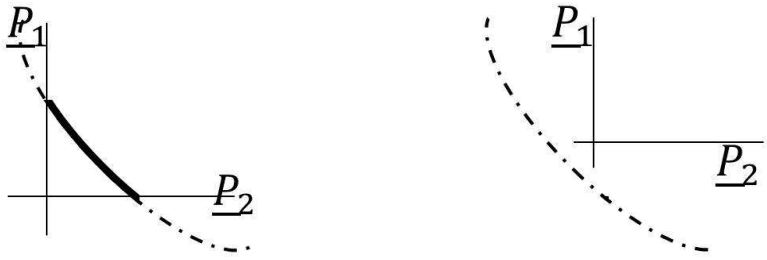


Figure 2.6: This figure illustrates the angle constraints in a distribution network. A pair (α, β) is assigned to each line in this figure, where α shows the angle between the two related buses under typical operating conditions (as given in the data) and β shows the limit from thermal constraints.



(a) Feasible (b) Infeasible

Figure 2.7: Either $\mathcal{P} = \mathcal{O}(\mathcal{P}) = \mathcal{O}(\text{conv}(\mathcal{P}))$ or the injection region is empty.

is made in Chapter 15 of [55], under which a practical optimization can be convexified (after approximating the power balance equations). However, our condition (2.19) is far less restrictive than (2.20). To understand the reason, note that the value $\frac{g_{12}}{b_{12}}$ is around 0.1 for a typical transmission line at the transmission level of the network [55] . Now, our condition allows for an angle difference as high as 80^o while the condition reported in [55] confines the angle to 6^o.

2.4.4 Tree Networks with All Active Power Constraints

In this section, we study general tree networks with local angle constraints and global bus power constraints. For every bus $i \in \mathcal{V}$, let $\overline{V_i}$ denote the fixed voltage magnitude $|V_i|$. Given an edge $(i,k) \in \mathcal{E}$, assume that the angle difference θ_{ik} belongs to the interval $[\underline{\theta}_{ik}, \overline{\theta}_{ik})]$, where $\underline{\theta}_{ik} \in [-\pi, 0]$ and $\overline{\theta}_{ki} \in [0, \pi]$. Define the angle-constrained flow region for the line (i,k) as

$$\mathcal{F}_{\theta ik} = \{(P_{ik}, P_{ki}) : \underline{\theta}_{ik} \leq \theta_{ik} \leq \overline{\theta}_{ik}, |V_i| = \overline{V}_i, |V_k| = \overline{V}_k\}$$

The angle-constrained injection region can be expressed as $\mathcal{P}_\theta = \mathbf{A}\mathcal{F}_\theta$, where $\mathcal{F}_\theta = \prod_{(i,k)\in\mathcal{E}} \mathcal{F}_{\theta ik}$. Following the insight from the last subsection, we make the following practical assumption

$$-\tan^{-1}(\frac{b_{ik}}{g_{ik}}) < \underline{\theta}_{ik} \leq \overline{\theta}_{ik} < \tan^{-1}(\frac{b_{ik}}{g_{ik}}), \quad \forall (i,k) \in \mathcal{E}. \tag{2.21}$$

This ensures that all points in the flow region of every line are Pareto optimal. As will be shown later, this assumption leads to the invertibility of the mapping from the injection region $\mathcal{P}_\theta$ to the flow region $\mathcal{F}_\theta$, or equivalently the uniqueness of the solution of every power flow problem.

Assume that the power injection P_i must be within the interval $[\underline{P}_i, \overline{P}_i]$ for every $i \in \mathcal{V}$. To account for these constraints, define the hyper-rectangle $\mathcal{P}_P = \{\mathbf{p} : \underline{\mathbf{p}} \leq \mathbf{p} \leq \overline{\mathbf{p}}\}$, where $\underline{\mathbf{p}} = (\underline{P}_1, \ldots, \underline{P}_n)$ and $\overline{\mathbf{p}} = (\overline{P}_1, \ldots, \overline{P}_n)$. The injection region $\mathcal{P}$ is then equal to $\mathcal{P}_\theta \cap \mathcal{P}_P$. In what follows, we present the main result of this section.

Theorem 4. *Suppose that $\mathcal{P}$ is a non-empty set. Under the assumption (2.21), the following statements hold:*

1. *For every injection vector $\mathbf{p} \in \mathcal{P}$, there exists a unique flow vector $\mathbf{f} \in \mathcal{F}$ such that $\mathbf{Af} = \mathbf{p}$.*
2. *$\mathcal{P} = \mathcal{O}(\mathcal{P})$.*
3. *$\mathcal{O}(\mathcal{P}) = \mathcal{O}(\text{conv}(\mathcal{P}))$.*

In order to prove this theorem, the next lemma is needed.

Lemma 5. *Under the assumptions of Theorem 4, the relation $\mathcal{O}(\mathcal{P}) = \mathcal{O}(\text{conv}(\mathcal{P}_\theta) \cap \mathcal{P}_P)$ holds.*

The proof of this lemma is provided in the appendix. Using this lemma, we prove Theorem 4 in the sequel.

Proof of Part 1: Given $\mathbf{p} \in \mathcal{P}$, consider an arbitrary leaf vertex k. Assume that i is the parent of bus k. Since k is a leaf, we have $P_{ki} = P_k$, and subsequently P_{ik} can be uniquely determined using the relation $\mathcal{F}_{jk} = \mathcal{O}(\mathcal{F}_{jk})$. One can continue this procedure for every leaf vertex and then go up the tree to determine the flow along each line in every direction.

Proof of Part 2: Since $\mathcal{P}$ is a subset of $\mathcal{P}_\theta$, it is enough to show that $\mathcal{P}_\theta = \mathcal{O}(\mathcal{P}_\theta)$. To prove this, the first observation is that $\mathcal{F}_\theta = \mathcal{O}(\mathcal{F}_\theta) = \mathcal{O}(\text{conv}(\mathcal{F}_\theta))$. Given a point $\mathbf{p} \in \mathcal{P}_\theta$, let $\mathbf{f} \in \mathcal{F}_\theta$ be the unique flow vector such that $\mathbf{Af} = \mathbf{p}$. There exist strictly positive numbers $\{c_{ik}, (i,k) \in \mathcal{E}\}$ such that $\mathbf{f}$ is the optimal solution to the following optimization problem

$$\mathbf{f} = \arg\min_{\tilde{\mathbf{f}} \in \mathcal{F}_\theta} \sum_{(i,k)\in\mathcal{E}} c_{ik}\tilde{P}_{ik}. \tag{2.22}$$

Since minimizing a strictly increasing function gives rise to a Pareto point, it is enough to show that there exists a set of positive constants $c_1, c_2, \ldots, c_n$ such that the optimal solution of the above optimization does not change if its objective function (2.22) is replaced by $\sum_{(i,k)\in\mathcal{E}} c_i\tilde{P}_{ik} = \sum_{i=1}^n c_i\tilde{P}_i$. Since $\mathcal{F}_\theta$ is a product space, we can multiply any pair (c_{ik}, c_{ki}) by a positive constant, and $\mathbf{f}$ still remains an optimal solution. Assume that the tree is rooted at 1. Let i be a leaf of the tree and consider the path from 1 to i. Without loss of generality, assume that the nodes on the path are labeled as $1, 2, \ldots, i$. By setting c_1 as c_{12}, one can define $c_2, \ldots, c_i$ according to the following recursion

$$c_k = c_{k-1}\frac{c_{k,k-1}}{c_{k-1,k}},$$

where k ranges from 2 to i. After defining $c_1, \ldots, c_i$, we remove all lines of the path 1–i from the network. This creates i disconnected subtrees of the network rooted at $1, \ldots, i$. For each of the subtrees with more than 1 node, one can repeat the above cost assignment procedure until $c_1, \ldots, c_n$ have all been constructed. This completes the proof. ■

Proof of Part 3: For notational simplicity, denote $\text{conv}(\mathcal{P}_\theta) \cap \mathcal{P}_P$ as $\mathcal{S}$. To prove this part, we use the relation

$$\mathcal{P} \subseteq \text{conv}(\mathcal{P}) \subseteq \mathcal{S} \tag{2.23}$$

and the result of Lemma 5, i.e.,

$$\mathcal{O}(\mathcal{P}) = \mathcal{O}(\mathcal{S}) \tag{2.24}$$

The first goal is to show the relation $\mathcal{O}(\mathcal{P}) \subseteq \mathcal{O}(\text{conv}(\mathcal{P}))$ by contradiction. Consider a vector $\mathbf{p} \in \mathcal{O}(\mathcal{P})$ such that $\mathbf{p} \notin \mathcal{O}(\text{conv}(\mathcal{P}))$. There exists a vector $\mathbf{p}' \in \mathcal{O}(\text{conv}(\mathcal{P}))$ such that $\mathbf{p}' \leq \mathbf{p}$ with strict inequality in at least one coordinate. Hence, it follows from (2.23) that $\mathbf{p}$ is not a Pareto point of $\mathcal{S}$, while it is a Pareto point of $\mathcal{P}$. This contradicts (2.24). To prove the converse statement $\mathcal{O}(\text{conv}(\mathcal{P})) \subseteq \mathcal{O}(\mathcal{P})$, consider a point $\mathbf{p} \in \mathcal{O}(\text{conv}(\mathcal{P}))$. In light of (2.23), $\mathbf{p}$ belongs to $\mathcal{S}$. If $\mathbf{p} \in \mathcal{O}(\mathcal{S})$, then $\mathbf{p} \in \mathcal{O}(\mathcal{P})$ due to (2.24). If $\mathbf{p} \notin \mathcal{O}(\mathcal{S})$, then there must exist a point $\mathbf{p}' \in \mathcal{O}(\mathcal{S}) = \mathcal{O}(\mathcal{P})$ such that $\mathbf{p}' \leq \mathbf{p}$ with strict inequality in at least one coordinate. This implies $\mathbf{p}' \in \mathcal{P}$ and consequently $\mathbf{p}' \in \text{conv}(\mathcal{P})$, which contradicts $\mathbf{p} \in \mathcal{O}(\text{conv}(\mathcal{P}))$. ■

2.4.5 Reactive Power Constraints

In the previous chapters, we have focused on the active powers. In this section, we show that reactive power flow can be thought as a rotation of active power flow, and therefore much of the argument for active power carries over.

The reactive power flows between bus i and k is given by

$$Q_{ik} = |V_i|^2 g_{ik} - |V_i||V_k| g_{ik} \sin\theta_{ik} - |V_i||V_k| b_{ik} \cos\theta_{ik}$$
$$Q_{ik} = |V_k|^2 g_{ik} + |V_i||V_k| g_{ik} \sin\theta_{ik} - |V_i||V_k| b_{ik} \cos\theta_{ik}.$$

Let $\mathcal{G}_{ik} \in \mathbb{R}^2$ denote the region that contain all the $[Q_{ik}\ Q_{ki}]^T$ that can be achieved by varying θ_{ik} between 0 and 2π; similar to $\mathcal{F}_{ik}$, $\mathcal{G}_{ik}$ is again a linear transformations of a circle. The center of $\mathcal{G}_{ik}$ is $[b_{ik}\ b_{ik}]^T$, its major axis is parallel to $[1\ 1]^T$ and has length b_{ik}, while its minor axis is parallel to $[1, -1]^T$ and has length g_{ik}. The active and reactive power injection regions are related by a linear invertible mapping, as shown in Fig. 2.8: $\mathcal{G}_{ik} = \mathbf{H}_{ik}\mathcal{F}_{ik}$, with

$$\mathbf{H}_{ik} = \frac{1}{2b_{ik}g_{ik}} \begin{bmatrix} b_{ik}^2 - g_{ik}^2 & b_{ik}^2 + g_{ik}^2 \\ b_{ik}^2 + g_{ik}^2 & b_{ik}^2 - g_{ik} \end{bmatrix}. \tag{2.25}$$

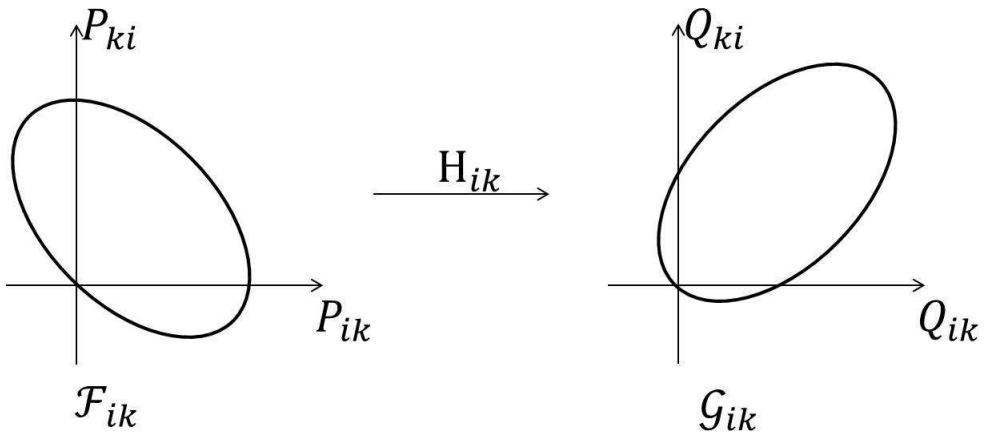


Figure 2.8: The active line flow region $\mathcal{F}_{ik}$, the reactive flow region $\mathcal{G}_{ik}$, and the linear transformation $\mathbf{H}_{ik}$ between them.

Again, we assume that the angle on a transmission line is limited between $\underline{\theta}_{ik}$ and $\overline{\theta}_{ik}$, where

$$-\tan^{-1}(b_{ik}/g_{ik}) < \underline{\theta}_{ik} < \overline{\theta}_{ik} < \overline{\theta}_{ik}.$$

Let $\mathcal{G}_{\theta,ik}$ be the angle-constrained reactive power flow region for the line i, k, and let $\mathcal{G}_\theta$ be the angle-constrained reactive power flow region. By (2.25), the reactive constraints can be thought as the constraints on the active flow region:

$$\mathcal{F}_Q = \{\mathbf{f} \in \mathbb{R}^{2(n-1)} : \underline{\mathbf{q}} \le \mathbf{A}(\mathbf{H}\mathbf{f}) \le \overline{\mathbf{q}}\}, \tag{2.26}$$

where $\mathbf{A}$ is the generalized edge-to-node matrix in (2.15); and $\mathbf{H}$ is the block diagonal matrix formed by stacking $\mathbf{H}_{ik}$ for each line (i, k). Following the same notation as before, we define the reactive power constrained active power injection region as

$$\mathcal{P}_Q = \mathbf{A}\mathcal{F}_Q. \tag{2.27}$$

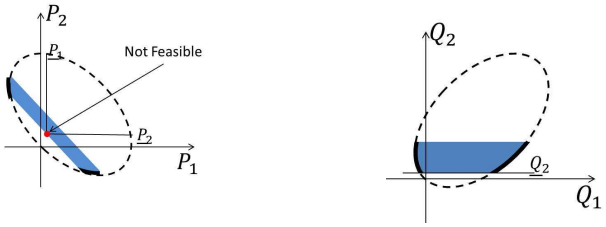


Figure 2.9: Active power injection region (left) and reactive power injection region (right) under reactive power injection lower bound.

The most general active injection region is given by

$$\mathcal{P} = \mathcal{P}_\theta \cap \mathcal{P}_P \cap \mathcal{P}_Q. \tag{2.28}$$

Theorem 6 below is the generalization of Theorem 4.

Theorem 6. *Suppose $\mathcal{P}$ is defined in* (2.28) *and it is non-empty. Furthermore, suppose the reactive power lower bounds satisfy*

$$\underline{Q}_i < \beta_i, \ \ i = 2, \ldots, n, \tag{2.29}$$

where $\beta_i = \sum_{k:k\in\mathcal{C}(i)} b_{ik} - g_{ik}\sin(\tilde{\theta}_{ik}) - b_{ik}\cos(\tilde{\theta}_{ik})$, where $\mathcal{C}(i)$ is the set of all neighbors of i. Then

$$\mathcal{O}(\mathcal{P}) = \mathcal{O}(\text{conv}(\mathcal{P})).$$

The proof of Theorem 6 is given in the appendix. We illustrate the concept using the two-bus injection region again. Because the reactive flow region is a rotation of the active flow region, constraints on reactive power becomes "slanted" when they are translated into active power constraints, as shown in Fig. 2.9.

Figure 2.9 shows the reason that the condition in (2.29) on the reactive power lower bounds are needed. Figure 2.9b gives the reactive injection region with a tight reactive lower bound on bus 2. Figure 2.9a shows the corresponding active power injection region. Observe that it is possible for the optimal solution of the relaxed problem to be not in the original injection region, while the original problem remains feasible. The condition $\underline{Q}_i < \beta_i$ rules out this phenomenon by ensuring that the reactive power lower bounds are never tight.

2.4.6 Variable Voltages

So far, we have assumed that all complex voltages in the network have fixed magnitudes. In this section, the results derived earlier will be extended to the case with variable voltage magnitudes under the assumption $\underline{P}_i = -\infty$ for every $i \in \mathcal{V}$. The goal is to study the

injection region after imposing the constraints

$$P_i \leq \overline{P}_i, \qquad i \in \mathcal{V} \tag{2.30a}$$

$$\theta_{ik} \in [\underline{\theta}_{ik}, \overline{\theta}_{ik}], \qquad (i,k) \in \mathcal{E} \tag{2.30b}$$

where $|\underline{\theta}_{ik}|, |\overline{\theta}_{ik}| < 90^\circ$. Note that the results to be developed next are valid even with explicit line flow constraints.

Given a bus $i \in \mathcal{V}$, let $\underline{V}_i$ and $\overline{V}_i$ denote the given lower and upper bounds on $|V_i|$. In vector notation, define $\underline{\mathbf{v}} = (\underline{V}_1, \ldots, \underline{V}_n)$ and $\overline{\mathbf{v}} = (\overline{V}_1, \ldots, \overline{V}_n)$. Given a vector $\tilde{\mathbf{v}} \in \mathbb{R}^+$, define $\mathcal{P}_\theta(\tilde{\mathbf{v}})$ as the angle-constrained injection region in the case when the voltage magnitudes are fixed according to $\tilde{\mathbf{v}}$, i.e. $|V_i| = \tilde{V}_i$ for $i = 1, ..., n$. Let $\mathcal{P}$ and $\mathcal{P}_\theta$ denote the regions for the case with variable voltage magnitudes. One can write $\mathcal{P} = \mathcal{P}_\theta \cap \mathcal{P}_P$, where

$$\mathcal{P}_\theta = \bigcup_{\underline{\mathbf{v}} \leq \tilde{\mathbf{v}} \leq \overline{\mathbf{v}}} \mathcal{P}_\theta(\tilde{\mathbf{v}}) \tag{2.31}$$

The problem of interest is to compute the convex hull of $\mathcal{P}_\theta$. However, the challenge is that the union operator does not commute with the convex hull operator in general (because the union of two convex sets may not be convex). In what follows, this issue will be addressed by exploiting the flow decomposition technique introduced in [53]. Let $\mathcal{H}_2^+$ denote the convex set of 2×2 positive semidefinite Hermitian matrices and $\mathcal{H}_n$ denote the set of all $n \times n$ Hermitian matrices. Given a matrix $\mathbf{W} \in \mathcal{H}_n$ together with an edge $(i,k) \in \mathcal{E}$, define:

- W_{ik}: (i,k) entry of $\mathbf{W}$.
- $\mathbf{W}_{ik}$: The 2×2 submatrix of $\mathbf{W}$ corresponding to the entries $(i,i), (i,k), (k,i), (k,k)$. The matrix $\mathbf{W}_{ik}$ is called an edge submatrix of $\mathbf{W}$.

Define also

$$\mathcal{H}_{ik}(\tilde{\mathbf{v}}) = \Big\{ \mathbf{W} \in \mathcal{H} : \mathbf{W}_{ik} \in \mathcal{H}_2^+, W_{ii} = \tilde{V}_i^2, W_{kk} = \tilde{V}_k^2, \\ \tan(\underline{\theta}_{ik}) \times \mathrm{Im}(W_{ik}) \leq \mathrm{Re}(W_{ik}), \\ \mathrm{Re}(W_{ik}) \leq \tan(\overline{\theta}_{ik}) \times \mathrm{Im}(W_{ik}) \Big\}$$

It can be shown that for every matrix $\mathbf{W} \in \mathcal{H}_{ik}(\tilde{\mathbf{v}})$ with the property $\mathrm{Rank}(\mathbf{W}_{ik}) = 1$, there exists an angle $\theta_{ik} \in [\underline{\theta}_{ik}, \overline{\theta}_{ik}]$ such that

$$\mathbf{W}_{ik} = \begin{bmatrix} \tilde{V}_i^2 & \tilde{V}_i \tilde{V}_k \angle \theta_{ik} \\ \tilde{V}_i \tilde{V}_k \angle \theta_{ki} & \tilde{V}_k^2 \end{bmatrix}$$

Thus,

$$\mathcal{F}_{\theta_{ik}}(\tilde{\mathbf{v}}) = \big\{ \mathrm{Re}(\mathrm{diag}(\mathbf{W}_{ik} \mathbf{Y}_{ik}^H)) : \mathbf{W} \in \mathcal{H}_{ik}(\tilde{\mathbf{v}}), \\ \mathrm{Rank}(\mathbf{W}_{ik}) = 1 \big\}$$

where $\mathbf{Y}_{ik} = \begin{bmatrix} y_{ik} & -y_{ik} \\ -y_{ik} & y_{ik} \end{bmatrix}$ (see [52, 53]). The flow region $\mathcal{F}_{\theta_{ik}}(\tilde{\mathbf{v}})$ can be naturally convexified by dropping its rank constraint. However, the convexified set may not be identical to $\mathrm{conv}(\mathcal{F}_{\theta_{ik}}(\tilde{\mathbf{v}}))$. We use the notation $\overline{\mathrm{conv}}(\mathcal{F}_{\theta_{ik}}(\tilde{\mathbf{v}}))$ for the convexified flow region, which is defined as

$$\overline{\mathrm{conv}}(\mathcal{F}_{\theta_{ik}}(\tilde{\mathbf{v}})) = \left\{ \mathrm{Re}(\mathrm{diag}(\mathbf{W}_{ik}\mathbf{Y}_{ik}^H)) : \mathbf{W} \in \mathcal{H}_{ik}(\tilde{\mathbf{v}}) \right\} \tag{2.32}$$

The following sets can also be defined in a natural way:

$$\overline{\mathrm{conv}}(\mathcal{F}_\theta(\tilde{\mathbf{v}})) = \prod_{(i,k)\in\mathcal{E}} \overline{\mathrm{conv}}(\mathcal{F}_{\theta_{ik}}(\tilde{\mathbf{v}})),$$
$$\overline{\mathrm{conv}}(\mathcal{P}_\theta(\tilde{\mathbf{v}})) = \mathbf{A}\overline{\mathrm{conv}}(\mathcal{F}_\theta(\tilde{\mathbf{v}}))$$

Note that $\mathrm{conv}(\cdot)$ and $\overline{\mathrm{conv}}(\cdot)$ were the same if the angle constraint (2.30b) did not exist.

Lemma 7. *Given a vector $\tilde{\mathbf{v}}$, the following relations hold:*

$$\mathcal{O}(\mathcal{P}_\theta(\tilde{\mathbf{v}})) = \mathcal{O}(\overline{\mathrm{conv}}(\mathcal{P}_\theta(\tilde{\mathbf{v}}))), \tag{2.33a}$$
$$\mathcal{O}(\mathcal{P}(\tilde{\mathbf{v}})) = \mathcal{O}(\overline{\mathrm{conv}}(\mathcal{P}_\theta(\tilde{\mathbf{v}})) \cap \mathcal{P}_P), \tag{2.33b}$$
$$\bigcup_{\underline{\mathbf{v}}\le\tilde{\mathbf{v}}\le\overline{\mathbf{v}}} \overline{\mathrm{conv}}(\mathcal{P}_\theta(\tilde{\mathbf{v}})) = \text{Convex set}. \tag{2.33c}$$

The proof is provided in the appendix.

As pointed out before Lemma 7, $\mathrm{conv}(\cdot)$ and $\overline{\mathrm{conv}}(\cdot)$ are equivalent if the angle constraint (2.30b) is ignored. In this case, it follows from (2.33c) and the relation

$$\mathcal{P}_\theta \subseteq \bigcup_{\underline{\mathbf{v}}\le\tilde{\mathbf{v}}\le\overline{\mathbf{v}}} \mathrm{conv}(\mathcal{P}_\theta(\tilde{\mathbf{v}})) \subseteq \mathrm{conv}(\mathcal{P}_\theta)$$

that $\mathrm{conv}(\mathcal{P}_\theta) = \bigcup_{\underline{\mathbf{v}}\le\tilde{\mathbf{v}}\le\overline{\mathbf{v}}} \mathrm{conv}(\mathcal{P}_\theta(\tilde{\mathbf{v}}))$. In other words, as long as there is no angle constraint, the convex hull operator commutes with the union operator when it is applied to (2.31). Motivated by this observation, define $\overline{\mathrm{conv}}(\mathcal{P}_\theta)$ as the convex set $\bigcup_{\underline{\mathbf{v}}\le\tilde{\mathbf{v}}\le\overline{\mathbf{v}}} \overline{\mathrm{conv}}(\mathcal{P}_\theta(\tilde{\mathbf{v}}))$. We present the main theorem of this section below.

Theorem 8. *For a tree network, $\mathcal{O}(\mathcal{P}) = \mathcal{O}(\mathrm{conv}(\mathcal{P})) = \mathcal{O}(\overline{\mathrm{conv}}(\mathcal{P}_\theta) \cap \mathcal{P}_P)$.*

Proof: Since

$$\mathcal{P} \subseteq \mathrm{conv}(\mathcal{P}) \subseteq \overline{\mathrm{conv}}(\mathcal{P}_\theta) \cap \mathcal{P}_P, \tag{2.34}$$

it suffices to prove that $\mathcal{O}(\mathcal{P}) = \mathcal{O}(\overline{\mathrm{conv}}(\mathcal{P}_\theta) \cap \mathcal{P}_P)$ (see part 3 of Theorem 4). First, we show that $\mathcal{O}(\overline{\mathrm{conv}}(\mathcal{P}_\theta) \cap \mathcal{P}_P) \subseteq \mathcal{O}(\mathcal{P})$. Consider a vector $\mathbf{p}$ in $\mathcal{O}(\overline{\mathrm{conv}}(\mathcal{P}_\theta) \cap \mathcal{P}_P)$. By the definition of $\overline{\mathrm{conv}}(\mathcal{P}_\theta)$, $\mathbf{p} \in \mathcal{O}(\overline{\mathrm{conv}}(\mathcal{P}_\theta(\tilde{\mathbf{v}})) \cap \mathcal{P}_P)$ for some $\tilde{\mathbf{v}}$. Hence, by Lemma 7, $\mathbf{p} \in \mathcal{O}(\mathcal{P}(\tilde{\mathbf{v}}))$ and consequently $\mathbf{p} \in \mathcal{P}$. Now, it follows from (2.34) and $\mathbf{p} \in \mathcal{O}(\overline{\mathrm{conv}}(\mathcal{P}_\theta) \cap \mathcal{P}_P)$ that $\mathbf{p} \in \mathcal{O}(\mathcal{P})$. The relation $\mathcal{O}(\mathcal{P}) \subseteq \mathcal{O}(\overline{\mathrm{conv}}(\mathcal{P}_\theta) \cap \mathcal{P}_P)$ can be proved in line with the proof of Lemma 5. ■

2.5 Numerical Algorithm for Convex Relaxation

The previous sections have focused on the geometric property of the power injection region, but to solve the OPF problem, we need an algebraic description of the convexification results. The relation $\mathcal{O}(\mathcal{P}) = \mathcal{O}(\text{conv}(\mathcal{P}))$ derived before states that the minimization of an increasing function over either the nonconvex set $\mathcal{P}$ or the convexified counterpart $\text{conv}(\mathcal{P})$ leads to the same solution. However, employing a numerical algorithm to minimize a function directly over $\text{conv}(\mathcal{P})$ is difficult due to the lack of efficient algebraic representations of $\text{conv}(\mathcal{P})$.

To address this issue, we decompose $\mathcal{P}$ as $\mathcal{P}_\theta \cap \mathcal{P}_P \cap \mathcal{P}_Q$ and then use the fact that $\text{conv}(\mathcal{P}_\theta)$, $\text{conv}(\mathcal{P}_P)$ and $\text{conv}(\mathcal{P}_Q)$ all have simple algebraic representations. Specifically, we use a rank relaxation technique. Recall the OPF problem in (2.4) has the general form

$$\begin{aligned}
\text{minimize } & f(P_1, P_2, \ldots, P_n) && \text{(2.35a)}\\
\text{subject to } & \underline{V}_i \le |V_i| \le \overline{V}_i && \text{(2.35b)}\\
& L_{ik} \le l_{ik} && \text{(2.35c)}\\
& P_{ik} \le \overline{P}_{ik} && \text{(2.35d)}\\
& \underline{P}_i \le P_i \le \overline{P}_i && \text{(2.35e)}\\
& \underline{Q}_i \le Q_i \le \overline{Q}_i && \text{(2.35f)}\\
& \mathbf{p} + j\mathbf{q} = \text{diag}(\mathbf{v}\mathbf{v}^H\mathbf{Y}^H). && \text{(2.35g)}
\end{aligned}$$

The resistive loss on the transmission line between buses i and k can be written as $L_{ik} = \mathbf{v}^H\mathbf{G}_{ik}\mathbf{v}$ where $\mathbf{G}_{ik}$ is a matrix with the (i,i)th entry and the (k,k)th entry being g_{ik}, and the (i,k)th entry and the (k,i)th entry being $-g_{ik}$ and all other entries being 0. The power flow from bus i to bus k can be written as $P_{ik} = \mathbf{v}^H\mathbf{A}_{ik}\mathbf{v}$, where $\mathbf{A}_{ik}$ is a matrix with (i,i)th entry g_{ik}, the (i,k)th entry $\frac{1}{2}(-g_{ik} - jb_{ik})$, the (k,i)th entry $\frac{1}{2}(-g_{ik} + jb_{ik})$ and all the other entries 0. Let $\mathbf{A}_i = \frac{1}{2}(\mathbf{E}_i\mathbf{Y} + \mathbf{Y}^H\mathbf{E}_i)$ where $\mathbf{E}_i$ is the diagonal matrix with 1 at the $(i,i)th$ entry and 0 everywhere else. Similarly let $\mathbf{B}_i = \frac{1}{2j}(\mathbf{Y}^H\mathbf{E}_i - \mathbf{E}_i\mathbf{Y})$. Then the powers injected at bus i is given by $P_i = \mathbf{v}^H\mathbf{A}_i\mathbf{v}$ and $Q_i = \mathbf{v}^H\mathbf{B}_i\mathbf{v}$. Therefore (2.4) can be equivalently written as

$$\begin{aligned}
\text{minimize } & f(\mathbf{p}) && \text{(2.36)}\\
\text{subject to } & \underline{V}_i \le |V_i| \le \overline{V}_i,\ \forall i\\
& \mathbf{v}^H\mathbf{G}_{ik}\mathbf{v} \le l_{ik}\ \forall i \sim k\\
& \mathbf{v}^H\mathbf{A}_{ik}\mathbf{v} \le \overline{P}_{ik}\ \forall i \sim k\\
& \underline{P}_i \le \mathbf{v}^H\mathbf{A}_i\mathbf{v} \le \overline{P}_i\\
& \underline{Q}_i \le \mathbf{v}^H\mathbf{B}_i\mathbf{v} \le \overline{Q}_i\\
& \mathbf{p} + j\mathbf{q} = \text{diag}(\mathbf{v}\mathbf{v}^H\mathbf{Y}^H).
\end{aligned}$$

To expose the potential non-convexity, we can equivalently write it as

$$\begin{aligned} \text{minimize } & f(\mathbf{p}) \\ \text{subject to } & \underline{V}_i^2 \le W_{ii} \le \overline{V}_i^2, \ \forall i \\ & \mathrm{Tr}(\mathbf{G}_{ik}\mathbf{W}) \le l_{ik} \ \forall i \sim k \\ & \mathrm{Tr}(\mathbf{A}_{ik}\mathbf{W}) \le \overline{P}_{ik} \ \forall i \sim k \\ & \underline{P}_i \le \mathrm{Tr}(\mathbf{A}_i\mathbf{W}) \le \overline{P}_i \\ & \underline{Q}_i \le \mathrm{Tr}(\mathbf{B}_i\mathbf{W}) \le \overline{Q}_i \\ & \mathbf{p} + j\mathbf{q} = \mathrm{diag}(\mathbf{W}\mathbf{Y}^H) \\ & \mathbf{W} \succcurlyeq 0 \\ & \mathrm{rank}(\mathbf{W}) = 1, \end{aligned} \tag{2.37}$$

where $\mathbf{W} = \mathbf{v}\mathbf{v}^H$ and the non-convexity enters as the rank 1 constraint on $\mathbf{W}$. Relaxing this rank 1 constraint and eliminating $\mathbf{p}$ and $\mathbf{q}$, we get

$$\begin{aligned} \text{minimize } & f(\mathrm{Tr}(\mathbf{A}_1\mathbf{W}), \ldots, \mathrm{Tr}(\mathbf{A}_n\mathbf{W})) \\ \text{subject to } & \underline{V}_i^2 \le W_{ii} \le \overline{V}_i^2, \ \forall i \\ & \mathrm{Tr}(\mathbf{G}_{ik}\mathbf{W}) \le l_{ik} \ \forall i \sim k \\ & \mathrm{Tr}(\mathbf{A}_{ik}\mathbf{W}) \le \overline{P}_{ik} \ \forall i \sim k \\ & \underline{P}_i \le \mathrm{Tr}(\mathbf{A}_i\mathbf{W}) \le \overline{P}_i, \\ & \underline{Q}_i \le \mathrm{Tr}(\mathbf{B}_i\mathbf{W}) \le \overline{Q}_i, \\ & \mathbf{W} \succcurlyeq 0. \end{aligned} \tag{2.38}$$

The optimization problem in (2.38) is convex, and if the conditions in Theorem 6 is satisfied, the optimal solution of (2.38) is rank 1, and the optimal value of (2.38) is the same as the optimal value of (2.4).

2.6 Network with Cycles

Ideally, one would like to generalize the results for trees to networks with cycles. However, this is difficult. In fact, there exist results stating the OPF problem for general power systems is NP-hard [56]. However, there is a large gap between tree networks and general electric networks, and in Appendix C section we provide partial progress in extending the results about tree networks.

2.7 DC Optimal Power Flow

In the previous sections we have focused on solving the AC power flow problem in distribution networks. Unfortunately, for transmission networks, it is known that the injection region

cannot be convex [56, 57]. On the other hand, the b/g ratio is very high ($b/g > 10$), and the DC approximation given below is often accurate enough.

The notations used in this thesis to describe DC power flow is somewhat new in the power system community and they are usually in graph theoretical terms.[6]

Given an electrical network with n buses and m transmission lines, we will think of it as a *directed graph* with the vertex set $\mathcal{V} = 1, \ldots, n$ and edge set $\mathcal{E} = \{(i,k) : i \sim k\}$. Note since the network is directed, the edge (i,k) means that it is originating from i and entering k. The set of flows on the edges is a vector of length m, indexed by the edge set, with $f_{ik} > 0$ meaning the flow is in the direction of the edge and $f_{ik} < 0$ if the flow is opposite to the direction. Let $\boldsymbol{\nabla}^T \in \mathbb{R}^{n \times m}$ be the mapping from branch flows to bus injections defined as

$$\boldsymbol{\nabla}^T(i,(k,l)) = \begin{cases} 1 & \text{if i=k} \\ -1 & \text{if i=l} \\ 0 & \text{otherwise} \end{cases}. \tag{2.39}$$

Let $\{b_{ik}\}$ be the set of susceptances, and let $\mathbf{b}$ be the vector of susceptances indexed by the edge set. Define

$$\mathbf{L} = \boldsymbol{\nabla}^T \operatorname{diag}(\mathbf{b}) \boldsymbol{\nabla}. \tag{2.40}$$

If the network is lossless and $\mathbf{Y}$ is purely imaginary, by a straightforward calculation,

$$\mathbf{Y} = -j\mathbf{L}. \tag{2.41}$$

The matrix $\mathbf{L}$ is called a Laplacian matrix, and the following lemma is true (see [58] for an in-depth discussion):

Lemma 9. $\mathbf{L}$ *is symmetric, positive semidefinite with rank* $n-1$. *Its null space is spanned by the* $\mathbf{1}$ *vector.*

In the DC approximation, the voltages are all assumed to be set at 1 per unit; the network is assumed to be lossless; and the angles are assumed to be small enough such that $\sin\theta \approx \theta$ and $\cos\theta \approx 1$ (that is, $\theta^2 \approx 0$). The power flow equation in (2.3) becomes

$$\begin{aligned} \mathbf{s} = \mathbf{p} + j\mathbf{q} &= \operatorname{diag}(\mathbf{v}\mathbf{v}^H \mathbf{Y}^H) \\ &\approx \operatorname{diag}((\mathbf{1} + j\boldsymbol{\theta})(\mathbf{1} + j\boldsymbol{\theta})^H \mathbf{Y}^H) \\ &= \operatorname{diag}((\mathbf{1}\mathbf{1}^T + j\boldsymbol{\theta}\mathbf{1}^T - j\mathbf{1}\boldsymbol{\theta}^T + \boldsymbol{\theta}\boldsymbol{\theta}^T)(j\mathbf{L}) \\ &\overset{(a)}{=} \operatorname{diag}((-j\mathbf{1}\boldsymbol{\theta}^T + \boldsymbol{\theta}\boldsymbol{\theta}^T)(j\mathbf{L}) \\ &\overset{(b)}{\approx} \operatorname{diag}((-j\mathbf{1}\boldsymbol{\theta}^T) j\mathbf{L} \\ &= \mathbf{L}\boldsymbol{\theta}, \end{aligned}$$

[6] We adopt graph-theoretical notations for two reasons: i) there is no established set of notations in power system engineering for stating our results; ii) some of the theorems are known in graph-theory but not in the form of DC power flow, so the notation can serve as a bridge between the two subjects.

where (a) follows from Lemma 9 and (b) follows from the assumptions that the angles are small. Notice only active power remains in the DC approximation. The DC-OPF is given by

$$\begin{aligned} &\text{minimize } f(\mathbf{p}) && \text{(2.42a)}\\ &\text{subject to } \underline{\mathbf{p}} \leq \mathbf{p} \leq \overline{\mathbf{p}} && \text{(2.42b)}\\ &\qquad |\mathbf{f}| \leq \mathbf{c} && \text{(2.42c)}\\ &\qquad \mathbf{p} = \mathbf{L}\boldsymbol{\theta} && \text{(2.42d)}\\ &\qquad \mathbf{f} = \operatorname{diag}(\mathbf{b})\boldsymbol{\nabla}\boldsymbol{\theta}. && \text{(2.42e)} \end{aligned}$$

It is convenient to eliminate $\boldsymbol{\theta}$ from (2.42) to write all variables in terms of the flows $\mathbf{f}$. To do this, we need to define the concept of an oriented sum over a cycle. Since the network is directed, a cycle in the network is directed as well. The oriented sum is the a sum of all flows in the cycle with respect to the orientation of the edges. An example of this is given in Fig. 2.10. A weighted oriented sum is an oriented sum where each flow is weighted by 1

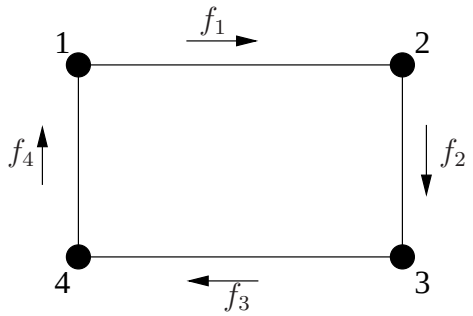


Figure 2.10: A four bus network. The weighted oriented sum is $f_1/b_{12}+f_2/b_{23}+f_3/b_{34}+f_4/b_{14}$.

over its admittance. We have the following lemma

Lemma 10. *There exist* $\boldsymbol{\theta} \in \mathbb{R}^n$ *such that* $\mathbf{f} = \operatorname{diag}(\mathbf{b})\boldsymbol{\nabla}\boldsymbol{\theta}$ *if and only if the weighted oriented sum of every cycle is* 0*.*

The proof of this lemma follows directly from the theory of fundamental cycles in graphs [59]. As an example, for the network in Fig. 2.10, the weighted oriented sum is

$$\begin{aligned} f_1/b_{12} + f_2/b_{23} + f_3/b_{34} + f_4/b_{14} &= \frac{b_{12}(\theta_1-\theta_2)}{b_{12}} + \frac{b_{23}(\theta_2-\theta_3)}{b_{23}} + \frac{b_{34}(\theta_3-\theta_4)}{b_{34}} + \frac{b_{14}(\theta_4-\theta_1)}{b_{14}}\\ &= \theta_1 - \theta_2 + \theta_2 - \theta_3 + \theta_3 - \theta_4 + \theta_4 - \theta_1\\ &= 0 \end{aligned}$$

We can represent constraint given by the oriented sums as a set of linear constraints on the flows, denoted by $\mathbf{Kf} = 0$. The DC OPF in (2.42) can be equivalently written as

$$\begin{aligned} \text{minimize } & f(\mathbf{p}) && \text{(2.43a)} \\ \text{subject to } & \underline{\mathbf{p}} \leq \mathbf{p} \leq \overline{\mathbf{p}} && \text{(2.43b)} \\ & |\mathbf{f}| \leq \mathbf{c} && \text{(2.43c)} \\ & \mathbf{p} = \boldsymbol{\nabla}^T \mathbf{f} && \text{(2.43d)} \\ & \boldsymbol{K}\mathbf{f} = 0. && \text{(2.43e)} \end{aligned}$$

Compared to AC-OPF, the feasible injection region for (2.43) is a polytope, thus (2.43) can be solved by standard optimization techniques if f is convex.

Chapter 3

Voltage Regulation

The objective of this section is to apply the result about distribution networks in Chapter 2 to the problem of voltage regulation with deep penetration of DERs; specifically, the focus is on the problem of mitigating voltage variability across the network due to fast (and uncontrolled) changes in the active generated or consumed by DERs. To this end, we rely on i) the use of the power electronics interfaces of the DERs to locally provide some limited amount of reactive power; and ii) to some extent, on the use of storage-capable DERs and DRRs to locally provide (or consume) some amount of active power. In other words, we have a limited ability to shape the active/reactive power injection profile. With respect to this, it is important to note that this ability to shape the active/reactive power injection profile, which in turn will allow us to regulate voltage across the network, it is intended to supplement the action of conventional voltage regulation devices (e.g., tap-changing under-load transformers, set voltage regulators, and fixed/switched capacitors).

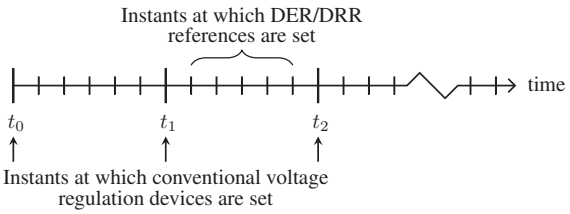


Figure 3.1: Time-scale separation between the instants in which the settings of conventional voltage regulation devices are decided, and the instants in which the reference of DER and DRRs are set.

In practice, in order to realize the ideas above, we envision a hierarchical control architecture in which there is a separation in the (slow) time-scale in which the settings of conventional voltage regulation devices are adjusted and the (fast) time-scale in which voltage regulation through active/reactive power injection shaping is accomplished. Then, given that fast (and uncontrolled) changes in the DERs active generation (consumption) might cause the voltage to deviate from this reference voltage, a second optimization is performed at regular intervals (e.g., every minute). The timeframe in which the settings of conven-

tional devices are decided and the reference setting of DERs/DRRs is graphically depicted in Fig. 3.1. The solution of this minute-by-minute optimization will provide the amount of active/reactive power that needs to be locally produced or consumed so as to track the voltage reference. In order words, the minute-by-minute optimization provides the reference values for the amount of active/reactive power to be collectively provided (or consumed) on each bus of the network within the next minute by reactive-power-capable and/or storage-capable DERs and DRRs. These reference values are then passed to the DERs and DRRs local controllers, which will adjust their output accordingly—note that the time-scale in which DER/DRR local controllers act (on the order of milliseconds (see, e.g., [60, 61]), is much faster than the minute-to-minute optimization.

3.1 Voltage Regulation via DERs/DRRs: Problem Formulation

As stated earlier, the focus of this paper is on developing mechanisms to mitigate voltage variability across the network due to fast (and uncontrolled) changes in the active generated or consumed by DERs; thus, subsequent developments only deal with the inter-hour minute-by-minute optimization mentioned above. As argued before, we assume that at the beginning of each hour, the settings of conventional voltage regulation devices are optimized, which in turns prescribes the value that each individual bus voltage can take to some voltage reference V_i^{ref}. In order to achieve the voltage regulation goal above, we rely on a limited ability to locally produce and/or consume some limited amount of active/reactive power, and cast the voltage regulation problem as an optimization program where the objective is to minimize network losses.

Let $L_{ik}(V_i, V_k) = P_{ik} + P_{ki}$; then, the total losses in the network are given by $L(\mathbf{v}) := \sum_{i,k:(i,k)\in\mathcal{E}} L_{ik}(V_i, V_k)$, and the voltage regulation problem can be formulated as

$$\min_{\mathbf{v}} \quad L(\mathbf{v}) \tag{3.1a}$$

$$\text{s.t.} \quad |V_i| = V_i^{ref}, \forall i \tag{3.1b}$$

$$\underline{P}_i \le P_i \le \overline{P}_i, \forall i \tag{3.1c}$$

$$\underline{Q}_i \le Q_i \le \overline{Q}_i, \forall i \tag{3.1d}$$

$$|P_{ik}| \le \overline{P}_{ik}, \ \forall i \sim k \tag{3.1e}$$

$$L_{ik}(V_i, V_k) \le \overline{L}_{ik}, \ \forall i \sim k \tag{3.1f}$$

$$P_i = \sum_{k\sim i} P_{ik} \tag{3.1g}$$

$$Q_i = \sum_{k\sim i} Q_{ik}. \tag{3.1h}$$

The constraints in (3.1b) capture the voltage regulation goal. The constraints in (3.1c) and (3.1d) describe the limited ability to control active/reactive power injections on each bus i; $\overline{P}_i$ ($\underline{P}_i$) and $\overline{Q}_i$ ($\underline{Q}_i$), denote the upper (lower) limits on the amount of active and reactive power that each bus i can provide, respectively. Without loss of generality and to ease the notations in subsequent development, hereafter we assume $V_i^{ref} = 1$ p.u. for all i. Note that active power and reactive power need not be controllable at every bus. If for a particular bus they are not controllable, in the optimization problem we set the bus active and/or reactive power upper and lower bounds to be equal, which essentially fixes the active and/or reactive on that bus.

By results from the previous section, the convex rank relaxation of (3.1) is

$$\min_{\mathbf{W}\succcurlyeq 0} \quad \sum_{i=1}^{n} \mathrm{Tr}(\mathbf{A}_i\mathbf{W}) \tag{3.2a}$$

$$\text{s.t.} \quad \mathbf{W}[i,i] = 1, \forall i \tag{3.2b}$$

$$\underline{P}_i \leq \mathrm{Tr}(\mathbf{A}_i\mathbf{W}) \leq \overline{P}_i, \forall i \tag{3.2c}$$

$$\underline{Q}_i \leq \mathrm{Tr}(\mathbf{B}_i\mathbf{W}) \leq \overline{Q}_i, \forall i \tag{3.2d}$$

$$\mathrm{Tr}(\mathbf{G}_{ik}\mathbf{W}) \leq \overline{L}_{ik}, \forall i \sim k \tag{3.2e}$$

$$|\mathrm{Tr}(\mathbf{A}_{ik}\mathbf{W})| \leq \overline{P}_{ik}, \forall i \sim k. \tag{3.2f}$$

3.2 A Distributed Algorithm for Solving the Convexified Problem

In Chapter 2, we showed that the SDP program in (3.2) is a convex relaxation of the voltage regulation problem in (3.1). Since the objective is to regulate the voltages in the presence of fast-changing power injection that, e.g., arise from renewable-based generation; the optimization problem needs to be solved no slower than the time-scale at which these injections significantly change. General-purpose SDP solvers scale poorly as the problem size increases [25]. Thus for large distribution networks with hundreds or thousands of buses, solving the SDP problem in a minute to sub-minute scale is challenging. Furthermore standard solvers for SDP problems are centralized; i.e., it is assumed that all the data defining the problem is available to a single processor. However the communication infrastructure in a distribution network may not be able to transmit all the data to a centralized location fast enough. By exploiting the tree structure of distribution networks, we propose a distributed algorithm to solve (3.2) that only requires communication between neighboring buses.

3.2.1 Algorithm Derivation

The proposed algorithm consists of two stages: *local optimization* and *consensus*. In the local optimization stage, each node solves its own local version of the problem. In the consensus

stage, neighboring nodes exchange Lagarangian multipliers obtained from the solutions to their corresponding local optimums, with the goal of equalizing the phase angle differences across a line from both of its ends.

Let $\mathcal{N}_i$ be the set of buses directly connected to bus i by transmission lines, together with bus i itself, i.e., $\mathcal{N}_i = \{k : k \sim i, \forall k\} \cup \{i\}$. For a $n \times n$ matrix $\mathbf{M}$, let $\mathbf{M}^{(i)}$ denote the $|\mathcal{N}_i| \times |\mathcal{N}_i|$ submatrix of $\mathbf{M}$ whose rows and columns are indexed according to $\mathcal{N}_i$. Similarly, for the $n \times 1$ vector $\mathbf{v}$, $\mathbf{v}^{(i)}$ is the corresponding $\mathcal{N}_i$-dimensional vector indexed by $\mathcal{N}_i$. We can rewrite (3.2) as

$$\min_{\mathbf{W}^{(1)},\ldots,\mathbf{W}^{(n)} \succcurlyeq 0} \quad \sum_{i=1}^{n} \mathrm{Tr}(\mathbf{A}^{(i)}\mathbf{W}^{(i)}) \tag{3.3a}$$

$$\text{s.t.} \quad \mathrm{diag}(\mathbf{W}^{(i)}) = \mathbf{v}^{(i)} \circ \mathbf{v}^{(i)}, \forall i \tag{3.3b}$$

$$\underline{P}_i \leq \mathrm{Tr}(\mathbf{A}^{(i)}\mathbf{W}^{(i)}) \leq \overline{P}_i, \forall i \tag{3.3c}$$

$$\underline{Q}_i \leq \mathrm{Tr}(\mathbf{B}^{(i)}\mathbf{W}^{(i)}) \leq \overline{Q}_i, \forall i \tag{3.3d}$$

$$|\mathrm{Tr}(\mathbf{A}_{ik}^{(i)}\mathbf{W}^{(i)})| \leq \overline{P}_{ik}, \forall (i,k) \in \mathcal{E} \tag{3.3e}$$

$$W_{ik}^{(i)} = W_{ik}^{(k)}, \forall (i,k) \in \mathcal{E}, \tag{3.3f}$$

$$W_{ki}^{(i)} = W_{ki}^{(k)}, \forall (i,k) \in \mathcal{E}, \tag{3.3g}$$

where $\circ$ is the Hadamard product. It is easy to verify that (3.3a), (3.3b), (3.3c), (3.3d), and (3.3e) are equivalent to (3.2a), (3.2b), (3.2c), (3.2d), and (3.2f), respectively, as $\mathbf{A}_i$ in (3.2c), $\mathbf{B}_i$ in (3.2d), and $\mathbf{A}_{ik}$ in (3.2f) have non-zero elements only at (i,i), (i,k), (k,i), $\forall k \sim i$. Since all $\mathcal{N}_i$'s are maximal cliques, $\mathbf{W} \succcurlyeq 0$ is tantamount to $\mathbf{W}^{(i)} \succcurlyeq 0, \forall i$ [62]. Constraints (3.3f) and (3.3g) are added to ensure that all $\mathbf{W}^{(i)}$'s coordinate to form $\mathbf{W}$; in other words, $\forall (i,k) \in \mathcal{E}$, the θ_{ik}'s computed from $\mathbf{W}^{(i)}$ and $\mathbf{W}^{(k)}$ should be the same.

Let λ_{ik} be the Lagrangian multiplier of (3.3f) for (i,k) and similarly λ_{ki} for (3.3g). By relaxing (3.3f) and (3.3g), the augmented objective function is

$$\sum_{i=1}^{n} \mathrm{Tr}(\mathbf{A}^{(i)}\mathbf{W}^{(i)}) + \sum_{(i,k)\in\mathcal{E}} [\lambda_{ik}(W_{ik}^{(i)} - W_{ik}^{(k)}) + \lambda_{ki}(W_{ki}^{(i)} - W_{ki}^{(k)})] \triangleq \sum_{i=1}^{n} \mathrm{Tr}(\tilde{\mathbf{A}}^{(i)}\mathbf{W}^{(i)}), \tag{3.4}$$

where $\tilde{\mathbf{A}}^{(i)}$ is also Hermitian, and its $(i,k)^{th}$ entry is i) $\tilde{A}_{ik}^{(i)} = A_{ik}^{(i)}$ if $i = k$, $\tilde{A}_{ik}^{(i)} = A_{ik}^{(i)} + \lambda_{ik}^{H}$ if $i < k$, and iii) $\tilde{A}_{ik}^{(i)} = A_{ik}^{(i)} - \lambda_{ik}^{H}$ if $i > k$. With (3.4), problem (3.3) can be divided into n

separable subproblems and the ith subproblem corresponds to bus i, defined as follows:

$$\min_{\mathbf{W}^{(i)} \succcurlyeq 0} \quad \mathrm{Tr}(\tilde{\mathbf{A}}^{(i)}\mathbf{W}^{(i)}) \tag{3.5a}$$

$$\text{s.t.} \quad \mathrm{diag}(\mathbf{W}^{(i)}) = \mathbf{v}^{(i)} \circ \mathbf{v}^{(i)} \tag{3.5b}$$

$$\underline{P}_i \leq \mathrm{Tr}(\mathbf{A}^{(i)}\mathbf{W}^{(i)}) \leq \overline{P}_i \tag{3.5c}$$

$$\underline{Q}_i \leq \mathrm{Tr}(\mathbf{B}^{(i)}\mathbf{W}^{(i)}) \leq \overline{Q}_i \tag{3.5d}$$

$$|\mathrm{Tr}(\mathbf{A}_{ik}^{(i)}\mathbf{W}^{(i)})| \leq \overline{P}_{ik}, \quad \forall k \sim i. \tag{3.5e}$$

We denote the feasible region described by (3.5b)–(3.5e) together with $\mathbf{W}^{(i)} \succcurlyeq 0$ of Subproblem i by $\mathcal{C}_i$. Define $g_i(\lambda_{ik}) \triangleq \inf_{\mathbf{W}^{(i)} \in \mathcal{C}_i}\{\mathrm{Tr}(\tilde{\mathbf{A}}^{(i)}\mathbf{W}^{(i)})\}$. The gradient of $-g_i$ at λ_{ik} is $W_{ik}^{(i)*}$, which is the (i,k)th element of the optimal $\mathbf{W}^{(i)*}$ of g_i determined by solving the ith subproblem (3.5). Similarly, that of $-g_k$ at λ_{ik} is $-W_{ik}^{(k)*}$. Therefore, the gradient of $-(g_i + g_k)$ is then $W_{ik}^{(i)*} - W_{ik}^{(k)*}$. Let $W_{ik}^{(i)}[t]$ and $W_{ik}^{(k)}[t]$ be $W_{ik}^{(i)*}$ and $W_{ik}^{(k)*}$ determined at time t, respectively. By gradient ascent, at time $t+1$, we update λ_{ik} by

$$\lambda_{ik}[t+1] = \lambda_{ik}[t] + \alpha[t](W_{ik}^{(i)}[t] - W_{ik}^{(k)}[t]), \tag{3.6}$$

where $\alpha[t] > 0$ and $\lambda_{ik}[t]$ are the step size and λ_{ik} at time t, respectively. The value of $\lambda_{ki}[t+1]$ can be directly computed from $\lambda_{ik}[t+1]$ as $\lambda_{ki} = \lambda_{ki}^H$. The Lagrangian multiplier λ_{ik} is only defined for the line (i,k) and the two buses at the ends of the edge, i.e., buses i and k, are required to manipulate λ_{ik}. The purpose of (3.6) is to make $W_{ik}^{(i)}$ and $W_{ik}^{(k)}$ as close to each other as possible with the help of λ_{ik}. Eq. (3.6) can be computed either by bus i or by bus k and it is independent of all other buses and edges. Whenever both the ith and kth subproblems have been computed and so $W_{ik}^{(i)}$ and $W_{ik}^{(k)}$ have been updated, then λ_{ik} can then be updated by using (3.6).

The optimization problem comprised of (3.4), together with all the constraints (3.5b)–(3.5e), imposed on the subproblems, is a dual problem of (3.3). When all λ_{ik}'s are optimal, $W_{ik}^{(i)}$ will be equal to $W_{ik}^{(k)}$ for all (i,k)'s and thus the duality gap is zero. Accordingly, we can construct the optimal $\mathbf{W}^*$ of problem from the values of the $W_{ik}^{(k)}$'s. Algorithm 1 can be seen as a dual decomposition algorithm, where the constraints on the consistence of line flows are dualized. Due to the convexity of (3.3), algorithm 1 converges to the optimal solution [63].

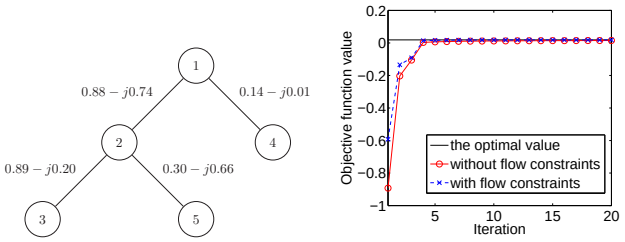
3.2.2 Feasibility

When the buses determine their own limits on active and reactive powers independently, an infeasible problem might result, i.e., an empty feasible region. When there exists a central authority having all the bus power information, we can check the feasibility easily. Otherwise, it is necessary for the buses to declare infeasibility.

One sufficient condition for infeasibility of the the problem is that there exists an infeasible subproblem (3.5) for any bus. If any bus finds an infeasible subproblem, it is sufficient

Given a n-bus network
1. **while** $|W_{ik}^{(i)} - W_{ik}^{(k)}| > \delta$ for any $(i,k) \in \mathcal{E}$ **do**
   2. **for** each bus i (in parallel) **do**
      3. Given $\lambda_{ik}, \forall k \sim i$, solve (3.5)
      4. Return $W_{ik}^{(i)}, \forall k$
   5. **end for**
   6.Given $W_{ik}^{(i)}$ and $W_{ik}^{(k)}$, update λ_{ik} with (3.6) (in parallel)
7. **end while**

Algorithm 1: Distributed Algorithm



(a) Network structure. (b) Convergence curve.

Figure 3.2: A five-bus example.

to say that the whole problem is infeasible. To proceed further, the bus with an infeasible subproblem should adjust its own active and reactive power limits so as to make the subproblem feasible. A necessary and sufficient condition for infeasibility is that $W_{ik}^{(i)}$ and $W_{ik}^{(k)}$ never match for some $(i,k) \in \mathcal{E}$ when Algorithm 1 evolves. If this happens on edge (i,k), either bus i or bus k or both constitute the infeasibility.

3.2.3 Numerical Performance Enhancements

Consider the five-bus network given in Fig. 3.2a. Assuming that all λ_{ik}'s are updated at the end of each iteration, the progress of Algorithm 1 (the curve without power flow constraints) and the target optimal objective value are shown in Fig. 3.2b. At iteration 20, when we sum the objective function values of all the subproblems, the sum still has around 20% difference to the optimal one. Even for a small network, it may take a long time for the algorithm to converge to the global optimal solution. Next, we provide some enhancements that improve the algorithm convergence speed.

Power Flow Constraints

Constraint (3.5e) means that the active power can flow in any direction on the edge (i,k) as long as its magnitude does not exceed the limit $\overline{P}_{ik}$. Assume that the global optimal solution

Table 3.1: Bus information of the five-bus example

Bus	$\overline{P}$	$\underline{P}$	$\overline{Q}$	$\underline{Q}$	$\overline{V}$
1	5.2844	-5.4692	5.5798	-5.7604	1.2247
2	-0.0648	-0.0988	0.5298	0	1.1509
3	-0.0423	-0.5828	0.6405	0	1.1103
4	-0.0334	-0.5155	0.2091	0	0.9762
5	-0.0226	-0.4329	0.3798	0	1.1400

$\mathbf{W}^*$ exists. Our decomposition allows us to compute W^*_{ik} separately by buses i and k, in which each bus determines its local version of W^*_{ik}, e.g., $W^{(i)}_{ik}$ for bus i. Then (3.6) brings both $W^{(i)}_{ik}$ and $W^{(k)}_{ik}$ towards W^*_{ik} by just equalizing $W^{(i)}_{ik}$ and $W^{(k)}_{ik}$. If the feasible regions $\mathcal{C}_i$ and $\mathcal{C}_k$ are smaller, it will be easier for (3.6) to reduce the discrepancy between $W^{(i)}_{ik}$ and $W^{(k)}_{ik}$.

The additional assumption we make is that all buses are net consumers of active power except the feeder; that is, $P_i \leq 0$ for $i = 2, 3, \ldots, n$. Even with deep penetration of renewable-based energy, it is unlikely that the power generated by these resources in a particular bus will compensate the load at that bus; let alone they will compensate for the power coming from the feeder. We now know that the active power must flow from buses i to k along the edge (i, k) with $i < k$, i.e., $P_{ik} \geq 0$. With this observation, we can re-write (3.5e) as

$$0 \leq P_{ik} = \mathrm{Tr}(\mathbf{A}^{(i)}_{ik}\mathbf{W}^{(i)}) \leq \overline{P}_{ik}, \tag{3.7}$$

$$-\overline{P}_{ik} \leq P_{ki} = \mathrm{Tr}(\mathbf{A}^{(k)}_{ik}\mathbf{W}^{(k)}) \leq 0, \tag{3.8}$$

from the perspectives of buses i and k, respectively. We can actually replace (3.5e) for (i, k) of Subproblem i by (3.7) and similarly (3.5e) for (i, k) of Subproblem k by (3.8). If we apply the same logic to all edges connecting to bus i, we can construct a smaller feasible region $\hat{\mathcal{C}}_i$ for Subproblem i. For the edge (i, k), the constructions of $\hat{\mathcal{C}}_i$ and $\hat{\mathcal{C}}_k$ can help $W^{(i)}_{ik}$ and $W^{(k)}_{ik}$ converge to W^*_{ik} faster.

With this modification, the progress of the algorithm for the five-bus example is also depicted in Fig. 3.2b, where we can see that the algorithm converges faster.

Feasible Solution Generation

When the algorithm converges, we have that

$$\mathrm{Tr}(\mathbf{A}^{(i)}\mathbf{W}^{(i)}) = \mathrm{Tr}(\tilde{\mathbf{A}}^{(i)}\mathbf{W}^{(i)}), \quad \forall i, \tag{3.9}$$

which holds when all its associated λ_{ik}'s are optimal; this is equivalent to have both of the following held:

$$\mathrm{Tr}(\mathbf{A}^{(i)}\mathbf{W}^{(i)}) = \mathrm{Tr}(\mathbf{A}^{(i)}\mathbf{W}^{(i)*}) \Leftrightarrow P_i = P^*_i, \quad \forall i, \tag{3.10}$$

$$\mathrm{Tr}(\mathbf{B}^{(i)}\mathbf{W}^{(i)}) = \mathrm{Tr}(\mathbf{B}^{(i)}\mathbf{W}^{(i)*}) \Leftrightarrow Q_i = Q^*_i, \quad \forall i. \tag{3.11}$$

In other words, Algorithm 1 tries to find the the optimal active and reactive power pair $[P_i^*, Q_i^*]^T$ for each bus i by manipulating λ_{ik}'s defined for the corresponding lines. The more lines are connected to a bus (i.e., the more λ_{ik}'s it involves), the more difficult is for (3.10) and (3.11) to hold. The $[P_i, Q_i]^T$ pair affects the $[P_k, Q_k]^T$ pair through λ_{ik}. Consider the situation where edge (i, k) is the only line connected to bus k except for bus i. When $[P_k, Q_k]^T$ becomes optimal, this helps bus i converge in the sense that this reduces the variations of $[P_i, Q_i]^T$ induced from bus k. When Algorithm 1 evolves, the $[P_k, Q_k]^T$ of leaf bus k converges first as a leaf bus has only one edge. Then, we have the buses connected to the leaf buses converged. We continue this process and finally go up to the feeder.

For any leaf node k, we have $P_k = P_{ki}$ and $Q_k = Q_{ki}$, where bus i is the only bus connected to bus k. When the algorithm evolves, we obtain $W_{ik}^{(k)*}$ from the solution of the kth subproblem (3.5) when $\mathrm{Tr}(\mathbf{A}^{(k)}\mathbf{W}^{(k)})$ and $\mathrm{Tr}(\mathbf{B}^{(k)}\mathbf{W}^{(k)})$ are equal to P_k^* and Q_k^*, respectively. Once we have fixed $W_{ik}^{(k)*}$, we can add the constraint $W_{ik}^{(i)} = W_{ik}^{(k)*}$ to the ith subproblem for bus i by passing a message containing the value of $W_{ik}^{(k)*}$ from bus k to bus i. In matrix form, this constraint is equivalent to $\mathrm{Tr}(\mathbf{C}^{(i)}\mathbf{W}^{(i)}) = \mathrm{Re}\{W_{ik}^{(k)*}\}$ and $\mathrm{Tr}(\mathbf{D}^{(i)}\mathbf{W}^{(i)}) = \mathrm{Im}\{W_{ik}^{(k)*}\}$, where $\mathbf{C}^{(i)} = (C_{lm}^{(i)}, l, m \in \mathcal{N}_i)$, with $C_{lm}^{(i)} = \frac{1}{2}$ if $l = i$ and $m = k$, $C_{lm}^{(i)} = \frac{1}{2}$ if $l = k$ and $m = i$, and $C_{lm}^{(i)} = 0$ otherwise; and $\mathbf{D}^{(i)} = (D_{lm}^{(i)}, l, m \in \mathcal{N}_i)$, with $D_{lm}^{(i)} = \frac{1}{2}j$ if $l = i$ and $m = k$, $D_{lm}^{(i)} = -\frac{1}{2}j$ if $l = k$ and $m = i$, and $D_{lm}^{(i)} = 0$ otherwise. In this case, we reduce the n-bus network into the $(n-1)$-bus one by removing bus k. When all other buses with positive active power flown from bus i (i.e. $\{l : l \sim i, l > i\}$) have been fixed and "removed", bus i becomes a leaf bus in the reduced network. This process continues until we find all $W_{lm}^*, \forall (l, m) \in \mathcal{E}$. The global solution $\mathbf{W}^*$ can be constructed from those W_{lm}^*'s. However, for any bus k, if we fix P_k and Q_k which is not optimal, these errors will make its connecting bus i being fixed afterwards result in incorrect P_i and Q_i, which are not optimal either. To achieve this, we observe $P_i[t]$ and $Q_i[t]$ for a certain time period and check if their variations are significant. Assume that we are at time t, for the active power, we can keep track of the previous T P_i's and the current $P_i[t]$, i.e. $[P_i[t-T], P_i[t-T+1], \ldots, P_i[t]]^T$. We can say that $P_i[t]$ has converged if its cumulative change is less than a certain threshold γ (e.g. 10^{-4}), i.e.,

$$\sum_{k=0}^{T-1} \frac{|P_i[t-T+k] - P_i[t-T+k+1]|}{|P_i[t-T+k]|} < \gamma; \tag{3.12}$$

with a similar condition for the reactive power.

Hot Start

The problem needs to be solved repeatedly; when there are changes to the active/reactive limits at any bus, we apply Algorithm 1 to the problem again. In each update, we usually have small variation between the new $\overline{P}_i$ and the previous ones and also for $\underline{P}_i$. Thus, in subsequent instances of the problem, the optimal angle difference across each line usually

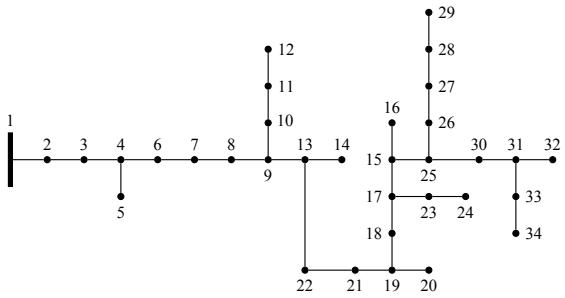


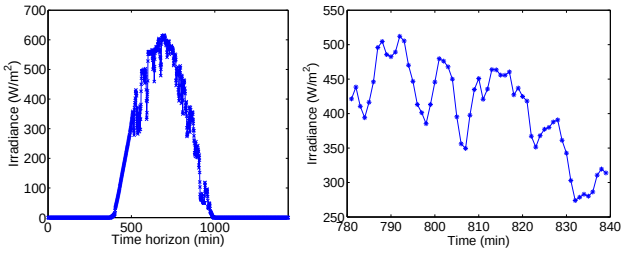
Figure 3.3: 34-bus system: electrical network graph. There are tap changing transformers between buses 7 and 8, buses 17 and 18, and buses 17 and 23.

does not vary significantly. Therefore, we can set $\lambda_{ik}[0]$ with the optimal λ_{ik}^* which can be determined from the previous optimal W_{ik}^*.

3.3 Case Studies

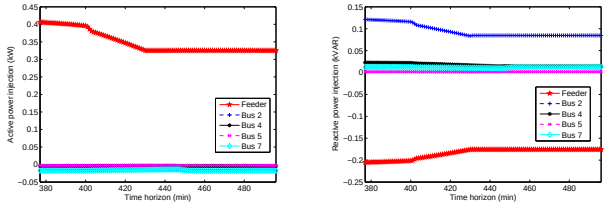
We test the performance of Algorithm 1 on the IEEE 34- and 123-bus test systems [50]; the data for these systems can be found in [50]. The topology for the 34-bus system is displayed in Fig. 3.3, while the topology for the 123-bus system is displayed in Fig. 2.1. All simulations were performed on a MacBookPro6,2, and each one was terminated when 300 iterations were reached.

Assume that, for both test systems, the nominal load on each bus i, which we denote by $\hat{P}_i$, is as specified by the datasets in [50]. Additionally, we assume that connected to each bus i, there are energy storage devices and PV-based electricity generation resources, which can supply active power, which we denote by P_i^{PV}, to the bus locally, i.e., their net effect is to reduce the load. If all P_i^{PV} is consumed locally, then the active power injection at bus i will be $\overline{P}_i = \hat{P}_i + P_i^{PV} \leq 0$. The computed optimal $P_i^* \in [\hat{P}_i, \overline{P}_i]$, $i = 2, \ldots, n$, will then be adjusted by controlling the amount of power from the PV devices which will be stored at the local storage device. Let $\hat{Q}_i$ be the nominal reactive power injection at bus i. Following [60], the power electronics interface of the PV installations is assumed to be able



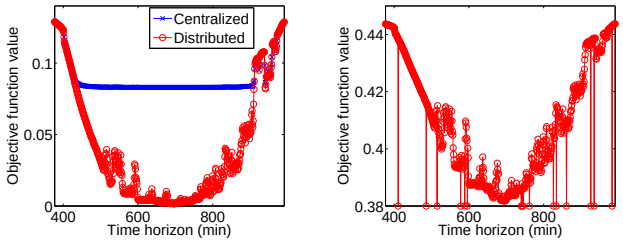
(a) Entire Time Horizon. (b) Minute 781 to 840.

Figure 3.4: Irradiance of a particular day in November 2011 [1].



(a) Active power injections. (b) Reactive power injections.

Figure 3.5: Active and reactive power injections at various buses in the 34-bus network.



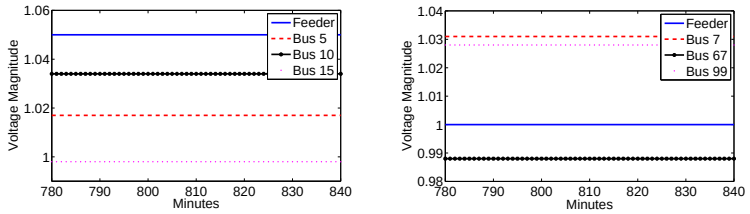
(a) 34-bus system (b) 123-bus system

Figure 3.6: Objective function values computed by the distributed algorithm.

to provide supply reactive power in a range that is sufficient to cancel the nominal reactive power [60]. Therefore, we assume that the reactive power can be adjusted in the ranges specified by i) $Q_i \in [0, 1.2\hat{Q}_i]$, if $\hat{Q}_i \geq 0$, and ii) $Q_i = [-1.2\hat{Q}_i, 0]$ otherwise.

Fits, we consider the one-minute resolution irradiance data in Fig. 3.4a, which correspond to a particular day in November 2011 collected at the University of Nevada [1]; the P_i^{PV}'s vary in accordance to the variation of this irradiance data. Assume that the PV systems connected to bus i can provide up to 20% of the nominal load $\hat{P}_i$ at that bus. Thus, the maximum P_i^{PV}, which is proportional to the respective $\hat{P}_i$, is different for different buses. As it can be seen in Fig. 3.4, since there is only radiation between the 377^{th} and 991^{th} minutes, for all numerical examples, we define a time horizon of $[377, 991]$, and execute Algorithm 1 every minute within this time horizon. Recall that Algorithm 1 requires inputs of Lagrangian multipliers as the starting points. In minute t, where $t \in [377, 991]$, the inputs to Algorithm 1 are the Lagrangian multipliers computed by Algorithm 1 at time $t-1$. Moreover each Lagrangian multiplier is only stored and manipulated by the two buses at the two ends of the corresponding transmission line. Initially, i.e., at $t = 377$, the Lagrangian multipliers are computed from the nominal system settings.

In order to check if the distributed algorithm can achieve the global optimum, we compare the objective function values computed by the distributed algorithm to those by the centralized solver; the results are plotted in Fig. 3.6. The active and reactive power injection at various buses in the 34-bus network are shown in Fig. 3.5a and Fig. 3.5b respectively.



(a) 34-Bus Network (b) 123-Bus Network

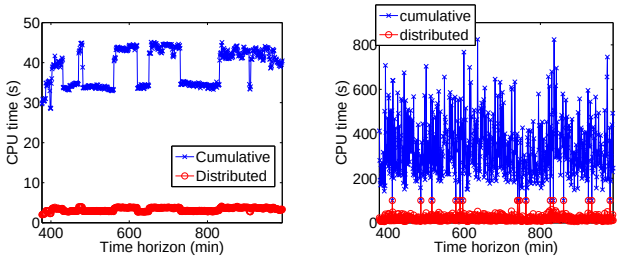
Figure 3.7: Voltage profile over time at representative buses. The proposed voltage regulation method is able to keep the voltages constant at their perspective references values.

As observed in Fig. 3.6, for the 34-bus system, we can see that the distributed algorithm converges to the optimum all the time, whereas for the 123-bus system convergence occurs most of the time. In Fig. 3.6b, the dropping lines correspond to the non-convergent cases where the convergence fails because the pre-defined 300 iterations allowed were exhausted. In the 34-bus and 123-bus systems, the centralized solver failed to solve the system due to convergence issues.

Figure 3.7 displays the voltage profile at various representative buses of the the 34- and 123-bus test system over a one-hour period with high variability in the the P_i^{PV}'s caused by the high-variability irradiance period displayed in Fig. 3.4b. This one-hour period corresponds to the portion of the daily irradiance profile in Fig 3.4a between the 781^{th} and 840^{th} minutes. For this simulation, the settings of the conventional voltage regulation devices are kept at the values given in [50], whereas the V_i^{ref}'s in (3.1b) result from the solution to the power flow equations for the nominal $\hat{P}_i$'s as specified in [50]. The fact that all the voltages displayed in Fig. 3.7 remain at their reference value illustrates the effectiveness of our proposed voltage regulation method to mitigate the effect of fast-varying power injections arising from PV systems.

Figure 3.8 shows the computational times corresponding to each test system; here we only consider the CPU time spent on the SDP solver and assume that communication overheads can be neglected. In our simulation, we implement the algorithm iteratively; in each iteration, we solve the subproblems sequentially. In Fig. 3.8, each subfigure contains two curves. One (distributed) is to sum the CPU times of the subproblems which need the longest CPU time in each iteration. In other words, we only consider the most demanding subproblem in each iteration and then sum the CPU times spent on these subproblems in all iterations. The average CPU computation time for the three cases are 1.28 s, 3.33 s, and 19.69 s, respectively, which are substantially shorter than the one-minute cycles considered.

Next, we show that the distributed algorithm is robust against random communication link failures. We model communication failures as packets drops. This means that, at a given iteration, the Lagrangian multiplier transmitted on any particular edge could be lost with probability p, independent of all other transmissions. Figure 3.9 shows the average time



(a) 34-bus system (b) 123-bus system

Figure 3 gorithm.

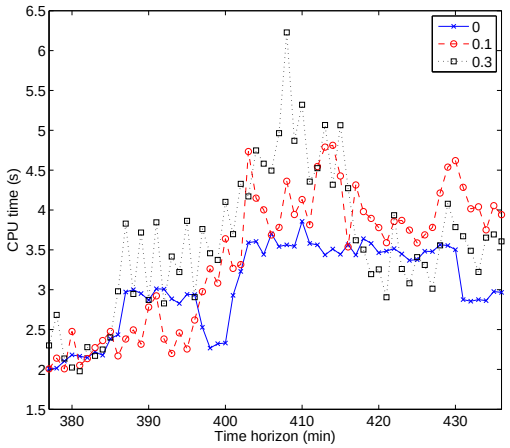


Figure 3.9: Time it takes for Algorithm 1 to converge under the presence of communication link failures.

of convergence needed over the day for the 34-bus network for $p = 0$, $p = 0.1$ and $p = 0.3$. Note that convergence is achieved at all times.

Chapter 4

Generator Dispatch

This chapter we focus on the problem of generator scheduling, or it is sometimes called, stochastic economic dispatch.

4.1 Problem Formulation

In current system operations, the generator is scheduled about 24 hours ahead of the actual time of electricity delivery [64]. This forward scheduling method is used because of the limited ramping capabilities on the generators. Figure 4.1 illustrate the scheduling process in detail.

Figure 4.1: The scheduling methodology used by system operators

.

In the US, most generator are scheduled through a market structure. Each generator and load serving entity (utility) submit bids to the system operator (SO), and the SO performs an double auction to clear this market. This process is the day-ahead (DA) market in Fig. 4.1. The outcome of the day-ahead market is a set of committed generators, with their scheduled generation levels, and a set of locational marginal prices. After this schedule is known, the SO will solve an optimal power flow problem with the set of committed generators and the forecasted load to determine if there is sufficient resources in the system. If there is an inadequacy of resources, the system operator would procure additional resources. This is the reliability assessment process in Fig. 4.1. Then at real-time, the actual delivery of power happens. To find the exact flows and generation at real-time, the SO solves another OPF problem with the actual demand.

The problem we focus on in this chapter is the reliability assessment process when there is significant wind power in the system. Before introducing wind, we first describe the current

reliability assessment problem in detail. Currently, only load is considered, and since load can be forecasted very accurately (within 1%), the SO performs the reliability assessment process by solving the following deterministic DC-OPF problem (see 2.7)

$$\text{minimize } \boldsymbol{\alpha}^T \mathbf{g} \tag{4.1a}$$

$$\text{subject to } \mathbf{g} - \boldsymbol{\nabla} \mathbf{f} \geq \hat{\mathbf{l}}, \tag{4.1b}$$

$$|\mathbf{f}| \leq \mathbf{c} \tag{4.1c}$$

$$\boldsymbol{K} \mathbf{f} = 0, \tag{4.1d}$$

where $\boldsymbol{\alpha} \in \mathbb{R}^n$ is the set of locational marginal prices at the buses; $\hat{\mathbf{l}}$ is the set of forecasted loads; and the optimization variables are $\mathbf{g}$ and $\mathbf{f}$, the generation levels and line flows, respectively. Since the optimization problem in (4.1) may not always be feasible, the power balance constraint in (4.1b) is typically written as a penalty term in the objective function, and (4.1) becomes

$$\text{minimize } \boldsymbol{\alpha}^T \mathbf{g} + \beta \mathbf{y}^+ \tag{4.2a}$$

$$\text{subject to } \mathbf{g} - \boldsymbol{\nabla} \mathbf{f} + \mathbf{y} = \hat{\mathbf{l}}, \tag{4.2b}$$

$$\underline{\mathbf{c}} \leq \mathbf{f} \leq \overline{\mathbf{c}} \tag{4.2c}$$

where β is a penalty associated with the cost of not meeting the load; $\mathbf{y}$ can be thought as a virtual generator that can absorb excessive power and has cost β if it is generating; and $(x)^+ = x$ if $x > 0$ is 0 otherwise.

With significant penetration of wind, the reliability assessment is no longer a deterministic problem. We assume that wind is taken as negative load, and define the net demand as

$$\mathbf{d} = \mathbf{l} - \mathbf{w}, \tag{4.3}$$

where $\mathbf{w}$ is the wind power. We use an additive stochastic model, where

$$\mathbf{d} = \hat{\mathbf{d}} + \mathbf{e}, \tag{4.4}$$

where $\mathbf{d}$ is the prediction and $\mathbf{e}$ is the prediction error. To define the stochastic version of (4.2), first consider the real-time OPF given by

$$J(\beta, \mathbf{g}) = \min_{\mathbf{g}^R} \beta^T (\mathbf{g}^R)^+ \tag{4.5a}$$

$$\text{subject to } \mathbf{g}^R - (\mathbf{d} - \mathbf{g}) - \boldsymbol{\nabla}^T \mathbf{f} = 0, \tag{4.5b}$$

$$\boldsymbol{K} \mathbf{f} = 0, \tag{4.5c}$$

$$|\mathbf{f}| \leq \mathbf{c}, \tag{4.5d}$$

where $\mathbf{g}$ represent the day-ahead decision that has already been made, the penalty is β and the realization of the random demand is known. The day-ahead stochastic RA problem is:

$$V^*(\hat{\mathbf{d}}) = \min_{\mathbf{g} \geq 0} \left\{ \boldsymbol{\alpha}^T \mathbf{g} + \mathbb{E}[J(\boldsymbol{\beta}, \mathbf{d} - \mathbf{g}) | \hat{\mathbf{d}}] \right\}, \tag{4.6}$$

where the expectation is taken with respect to the distribution of $\mathbf{d}$ conditional on the forecast $\hat{\mathbf{d}}$. The constraint $\mathbf{g} \geq 0$ limits the day ahead decisions to purchasing generation power only. Additionally, $\mathbf{g}$ is function of the forecast $\hat{\mathbf{d}}$ and the error distribution. We call the optimal solution to (4.6) the *risk limiting dispatch*.

There are two reasons why we focus on the intermediate RA problem rather than directly at the day-ahead market.

1. The day-ahead market is purely financial. That is, the bidders in the market can place arbitrary bids, and these bids may not correspond to any physical loads in the system. On the other hand, the system operator's role is simply to perform the double auction. Without changing the market structure significantly, there is no place to take the randomness into account. In contrast, the reliability assessment is performed by the system operator with actual load forecast and must take wind into account.

2. The day-ahead market must make commitment decisions and this results in a integer-valued optimization problem. In the reliability assessment, most of commitment decisions are already made by the day-ahead market, therefore only the generation levels need to be decided.

4.1.1 Integration Cost and Price of Uncertainty

A fundamental quantity of interest is the impact of uncertainty in the cost of dispatch. We call this quantity if integration cost [65], which is defined the *difference between the expected cost of the procedure and the expected cost of a dispatch clairvoyant of the realization of* $\mathbf{d}$. The clairvoyant dispatch can allocate all the required power in the day ahead by solving the deterministic OPF $V_C^*(\mathbf{d}) = J(\alpha, \mathbf{d})$. The integration cost for a realization of the information set $\hat{\mathbf{d}}$ is given by

$$C_I(\hat{\mathbf{d}}) = V^*(\hat{\mathbf{d}}) - \mathbb{E}[V_C^*(\hat{\mathbf{d}} + \mathbf{e})|\hat{\mathbf{d}}]. \tag{4.7}$$

An important question is regarding the sensitivity of this cost to the forecast error standard deviation σ_e when the *best possible dispatch* is utilized. If C_I is a linear function of σ_e, so $C_I = p\sigma_e$, then p is the *price of uncertainty*, a fundamental limit faced by *any* dispatch procedure. In this paper we show the existence of this price, and determine it for various scenarios.

4.1.2 Small-σ Assumption

An important consideration is the order of magnitude of the error standard deviation σ_e compared to the entries in the average net load vector μ and the transmission line capacities. Standard deviation of day ahead load forecasts σ_L are $1\% - -2\%$ of the expected load μ_L. Wind error forecasts are more severe, and error standard deviations σ_W of 30% of rated capacity μ_W have been observed. High wind penetration scenarios have about 30% of total

load being generated by wind, and therefore the total error would be about $0.01 + 0.3 * 0.3 = 10\%$ of total load.

In contrast to the financial situation, a relative forecast error of 10% would not change the overall physical operating characteristic of the network. More precisely, suppose we calculate the deterministic dispatch based on the forecast values $\hat{\mathbf{d}}$ and find bus i would be generating power in the first stage. Then with high probability, bus i would still be generating power in the two stage dispatch problem. Also, the network congestion pattern under the deterministic dispatch and the two-stage dispatch should not be drastically different. We formalizes these observations in later sections.

We denote above scenarios *small-σ* scenarios. In theoretic arguments this usually implies studying scaling regimes such that the variance of the prediction error $\sigma \to 0$. For simplicity of exposition, we avoid such limits in the relevant points in the analysis, but the formal arguments follow easily from the stated theorem. The important observation is that forecast values are very useful in determining the *qualitative* behaviour of the network.

4.2 Uncongested Network

This section reviews the risk limiting dispatch control for a two stage uncongested network [66, 67], and analyzes the price of uncertainty in this scenario.

4.2.1 Risk Limiting Dispatch

Since we only consider a single bus, all variables are scalar. Equivalently, the single bus network can be thought as an n-bus network without congestion. In this case, the constraint region in (4.5) reduces to net supply must equal net demand, and the RT-OPF becomes

$$\begin{aligned} J^*(\beta, d-g) = &\min \ \beta(g^R)^+ \\ &\text{s.t. } g^R + g - d = 0 \\ = &(d-g)^+. \end{aligned}$$

The DA-SPF in Eq. (4.6) can then be reduced to

$$g^* = \arg\min_g \alpha g + \beta\, \mathbb{E}[(d-g)^+ | \hat{d}] \tag{4.8a}$$

$$\text{s.t. } g \geq 0. \tag{4.8b}$$

RLD can then be derived as follows. Consider the unconstrained optimization problem

$$\min_g \ \alpha g + \beta\, \mathbb{E}[(d-g)^+ | \hat{d}]. \tag{4.9}$$

Taking the subgradient with respect to g gives the optimality condition

$$\begin{aligned} 0 &= \alpha - \beta\, \mathbb{E}[\mathbf{1}(d - g > 0)|\hat{d}] \\ &= \alpha - \beta\, \mathbb{E}[\mathbf{1}(\hat{d} + e - g > 0)|\hat{d}] \\ &= \alpha - \beta \Pr(e > g - \hat{d}|\hat{d}), \end{aligned}$$

rearranging gives

$$\Pr(e > g - \hat{d}|\hat{d}) = Q(g - \hat{d}) = \frac{\alpha}{\beta}, \tag{4.10}$$

the optimal control g is given by inverting (4.10)

$$g = \hat{d} + Q^{-1}(\frac{\alpha}{\beta}). \tag{4.11}$$

Note it is possible that $g < 0$, it can be shown that the risk limiting dispatch g^* (optimal solution to the constrained problem in (4.8)) is given by thresholding

$$g^* = g^+ = [\hat{d} + Q^{-1}(\frac{\alpha}{\beta})]^+. \tag{4.12}$$

4.2.2 Price of Uncertainty

Since most power systems would not have 100% penetration in the near future, we assume that the net demand d, and its prediction $\hat{d}$, are positive. Then first we would show the price of uncertainty exists (i.e. the integration cost is linear in σ_e), and then calculate its value.

Theorem 11. *Suppose $d > 0$. Then $C(\hat{d})$ defined in (4.7) is linear and can be written as*

$$C(\hat{d}) = \sigma_e p, \tag{4.13}$$

where σ_e is the standard deviation of the error e and $p = \beta\phi(Q^{-1}(\frac{\alpha}{\beta}))$ ($\phi(\cdot)$ is the standard Gaussian density and $Q(\cdot)$ is the complimentary Gaussian cumulative density function).

Theorem 11 relies on the observation that if net demand is positive ($d > 0$), then it is always beneficial to purchase energy in the day ahead as the energy price is higher in real-time, so the optimal schedule must be positive $g^* > 0$. The positivity constraint in the simplified DA-SPF (Eq. (4.8)) is redundant, and the cost of uncertainty (Eq. (4.7)) can be

computed as

$$
\begin{aligned}
C(\hat{d}) &= V^*(\hat{d}) - V_C(\hat{d}) && \text{(4.14a)}\\
&= \min_g \ \alpha g + \beta\,\mathbb{E}[(d-g)^+|\hat{d}] - \alpha\,\mathbb{E}[d^+|\hat{d}] && \text{(4.14b)}\\
&\overset{(a)}{=} \min_g \ \alpha g + \beta\,\mathbb{E}[(d-g)^+|\hat{d}] - \alpha\,\mathbb{E}[d|\hat{d}] && \text{(4.14c)}\\
&= \min_g \ \alpha g + \beta\,\mathbb{E}[(\hat{d}+e-g)^+|\hat{d}] - \alpha\,\mathbb{E}[d|\hat{d}] && \text{(4.14d)}\\
&\overset{(b)}{=} \min_\Delta \ \alpha(\hat{d}+\Delta) + \beta\,\mathbb{E}[(e-\Delta)^+|\hat{d}] - \alpha\,\mathbb{E}[\hat{d}+e|\hat{d}] && \text{(4.14e)}\\
&= \min_\Delta \ \alpha\Delta + \beta\,\mathbb{E}[(e-\Delta)^+|\hat{d}] && \text{(4.14f)}\\
&\overset{(c)}{=} \sigma_e\{\min_{\Delta'} \ \alpha\Delta' + \beta\,\mathbb{E}[(z-\Delta')^+|\hat{d}]\} && \text{(4.14g)}\\
&= \sigma_e p, && \text{(4.14h)}
\end{aligned}
$$

where (a) follows from the assumption $d > 0$, (b) follows from setting $g = \hat{d} + \Delta$, (c) follows from changes from variables where $\Delta' = \Delta/\sigma_e$ and $z = e/\sigma_e$. p only depends on the standardized random variable z with unit variance and thus is independent of the forecast error standard deviation σ_e.

The structure of the optimal control was not used in the previous derivation, so the existence of the price of uncertainty p was demonstrated without calculating its value, which can be numerically calculated in close form as below. Using the same technique as solving (4.9) the optimal Δ' in (4.14g) is given by $Q^{-1}(\frac{\alpha}{\beta})$. The price of uncertainty p becomes

$$
\begin{aligned}
p &= \alpha Q^{-1}(\frac{\alpha}{\beta}) + \beta\,\mathbb{E}[(z - Q^{-1}(\frac{\alpha}{\beta}))^+] && \text{(4.15a)}\\
&= \alpha Q^{-1}(\frac{\alpha}{\beta}) + \beta \int_{Q^{-1}(\frac{\alpha}{\beta})}^{\infty} (z - Q^{-1}(\frac{\alpha}{\beta}))\phi(z)dz && \text{(4.15b)}\\
&= \alpha Q^{-1}(\frac{\alpha}{\beta}) + \beta(-\frac{\alpha}{\beta}Q^{-1}(\frac{\alpha}{\beta}) + \phi(Q^{-1}(\frac{\alpha}{\beta}))) && \text{(4.15c)}\\
&= \beta\phi(Q^{-1}(\frac{\alpha}{\beta})). && \text{(4.15d)}
\end{aligned}
$$

Figure 4.2 plots the price of uncertainty for different values of α/β with β set to be 1. Somewhat surprisingly, p is not monotonic in α/β and it goes to 0 as α/β approaches 0 or α/β approaches 1. Intuitively, when α/β is small, the day ahead cost is very low, and the SO can purchase sufficient amounts of energy to absorb the prediction error. In contrast, when α/β is close to 1, the day ahead and real-time costs are similar, so the SO waits until real-time to balance the system once the net load realization is completely known.

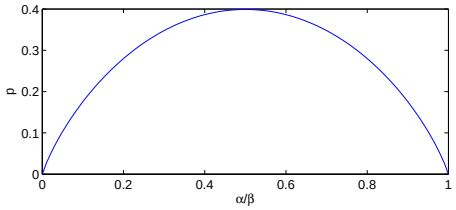


Figure 4.2: The price of uncertainty for different ratios of α/β.

4.2.3 Extremely High Penetration

In some networks renewable power may have a penetration level of more than 100%, violating the small-σ assumption. For example, in a microgrid where wind or solar energy is abundant, the net demand could become negative. In this case, the cost of uncertainty is no longer linear in the standard deviation of the prediction error and in general cannot be computed in closed form. Figure 4.3 plots the cost of uncertainty of a system where $\hat{d} = -1$ per unit, and $\alpha/\beta = 1/3$. The cost of uncertainty is low when σ_e is small since no additional energy

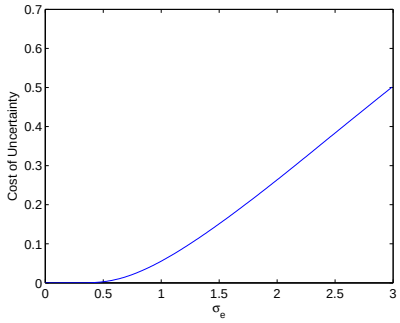


Figure 4.3: The cost of uncertainty for α/β=1/3, $\hat{d} = -1$.

is needed to balance the system, and as σ_e increases, the cost becomes linear with the slope $\phi(Q^{-1}(1/3))$.

4.3 Congested Networks

The RT-OPF in N-RLD for n-bus networks does not admit an analytical solution as in the single bus case, significantly increasing the complexity of the full dispatch. In particular, it is difficult to obtain the day ahead dispatch $\mathbf{g}$ in closed form. Moreover, the cost of uncertainty can be a complicated function of the information set Y and the network capacities $\mathbf{c}$. These quantities can be numerically computed resorting to a Monte Carlo approach, but the computational challenges are formidable due to the high dimensionality of the problem.

Instead, the small-σ assumption from Sec. 4.1.2 can be explored to obtain a simple and interpretable dispatch. Since the prediction error is a small percentage of the net load, the change in flows caused by that error is also a small percentage, we assume the prediction error is small compared to both $\hat{\mathbf{d}}$ and $\mathbf{c}$. Under the small-σ assumption, the *qualitative* or structural behavior of the power system predicted in the day-ahead from the forecast $\hat{\mathbf{d}}$ will not differ from its realization in real-time after observing $\mathbf{d}$. If we expect to purchase power at a bus in the day-ahead, then after real-time, we do not expect power to be shed in that bus. If a transmission line is expected to be congested in a certain direction in the day-ahead, then the direction of congestion would not be reversed at real time. Since qualitative features are consistent with the forecast, a deterministic OPF based on the day-ahead price $\boldsymbol{\alpha}$ and the net load forecast $\hat{\mathbf{d}}$ will predict congested lines, congestion directions and buses where energy is purchased correctly. This deterministic OPF is denominated *Nominal Day-Ahead OPF* (NDA-OPF):

$$J(\boldsymbol{\alpha}, \hat{\mathbf{d}}) = \min_{\mathbf{g}} \boldsymbol{\alpha}^T (\mathbf{g})^+ \tag{4.16a}$$

$$\text{subject to } \mathbf{g} - \hat{\mathbf{d}} \in \mathcal{P}. \tag{4.16b}$$

In stochastic control terms, NDA-OPF solves the certainty equivalent control problem for N-RLD (4.6) [68, 69], by replacing the random quantity $\mathbf{d}$ by the deterministic quantity $\hat{\mathbf{d}}$ and solving the optimization problem. Denote the generation schedule from NDA-OPF by $\overline{\mathbf{g}}$.

The day ahead schedule $\mathbf{g}$ in the DA-SPF (Eq. (4.6)) can be decomposed as the nominal dispatch added to a perturbation $\mathbf{g} = (\overline{g} + \boldsymbol{\Delta})^+$ where $\boldsymbol{\Delta} \in \mathbb{R}^n$ is the perturbation. The optimal schedule is determined by computing $\boldsymbol{\Delta}$. Perturbations are expected to be small since the uncertainty is small, so the *perturbed DA-SPF* can be significantly simplified. The simplification relies on three key observations for small-σ forecast error:

1. If $\overline{g}_i < 0$, bus i is treated as a source of unlimited energy, since it is shedding energy in the nominal problem (NDA-OPF). If $\overline{g}_i > 0$, then the perturbed dispatch Δ_i is not constrained to be positive.

2. If the line between buses i and k are not congested, then it is not congested in perturbed DA-SPF.

3. If the line between buses i and k is congested from i to k, then it would not become congested from k to i in the perturbed DA-SPF.

Going forward, we assume these observations to hold. This is called the **small-$\boldsymbol{\sigma}$** assumption. We propose the two step algorithm in Algorithm 1.

Algorithm 1: Procedure to solve Network Risk Limiting Dispatch

Step 1 (NDA-OPF): Solve the nominal problem in Eq. (4.16) using forecast net load and day ahead prices to obtain the nominal schedule $\overline{\mathbf{g}}$ and nominal line flows $\overline{\mathbf{f}}$.

Step 2 (Perturbed DA-SPF): Solve the DA-SPF (Eq. (4.6)) for the optimal perturbation

$\boldsymbol{\Delta}$ by substituting $\mathbf{g} = \overline{\mathbf{g}} + \sigma_e \boldsymbol{\Delta}^*$ and appropriately normalizing and reducing the problem using Observations $(1) - (3)$ as

$$\boldsymbol{\Delta}^* = \arg\min_{\boldsymbol{\Delta}} \boldsymbol{\alpha}^T \boldsymbol{\Delta} + \mathbb{E}[\tilde{J}(\boldsymbol{\beta}, \mathbf{e}) | \hat{\mathbf{d}}] \tag{4.17a}$$

$$\text{subject to } \Delta_i = 0 \text{ if } \overline{g}_i < 0, \tag{4.17b}$$

$$\Delta_i > 0 \text{ if } \overline{g}_i = 0, \tag{4.17c}$$

where

$$\tilde{J}(\boldsymbol{\beta}, \mathbf{e}) = \min \tilde{\boldsymbol{\beta}}^T (\mathbf{y})^+ \tag{4.18a}$$

$$\text{s.t. } \mathbf{y} - \mathbf{e} - \nabla^T \mathbf{f} = 0 \tag{4.18b}$$

$$\mathbf{K}\mathbf{f} = 0 \tag{4.18c}$$

$$f_{ik} < 0 \text{ if } \overline{f}_{ik} = c_{ik}, \tag{4.18d}$$

and $\tilde{\beta}_i = \beta_i$ if $\overline{g}_i \geq 0$ and $\tilde{\beta}_i = 0$ otherwise. The optimal DA-SPF dispatch is then given by $\mathbf{g} = (\overline{\mathbf{g}} + \sigma_e \boldsymbol{\Delta})^+$.

At first glance, (4.17) seems to be no simpler than the original problem in (4.6). However, note that the network capacity constraints (4.18d) only include the lines that are congested in the nominal problem. In essence, (4.17) balances a 'left-over' network from solving the nominal problem, and (4.18d) states that if a line is congested in the nominal problem, no more energy is allowed to flow along the direction of congestion.

The next subsection explores the normalization and reduction process to define the Perturbed DA-SPF for two bus and three bus networks. We show the perturbation $\boldsymbol{\Delta}$ is the solution to a set of deterministic equilibrium equations. Then the problem of an arbitrary network with n buses and a single congestion link is studied and we show the general reduction procedure results in an optimal dispatch control under the small-σ assumption.

4.3.1 Two Bus Network

Consider the two bus network in Fig. 4.4. For this network, the day ahead dispatch is a

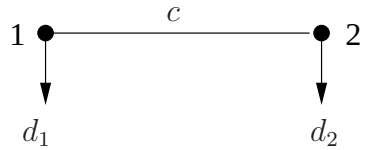


Figure 4.4: A two-bus network where c is the capacity of the line.

vector $\mathbf{g} = [g_1 \ g_2]^T$ of the scheduled generation at each bus. The real-time balancing of the network requires solving an OPF where the injection region is two dimensional. The

RT-OPF becomes

$$J(\boldsymbol{\beta}, \mathbf{d} - \mathbf{g}) = \min_{\mathbf{g}^R} \boldsymbol{\beta}^T (\mathbf{g}^R)^+ \tag{4.19a}$$

$$\text{subject to } g_1^R + g_1 - d_1 - f = 0 \tag{4.19b}$$

$$g_2^R + g_2 - d_2 + f = 0 \tag{4.19c}$$

$$|f| < c, \tag{4.19d}$$

where f is the amount of power flowing from bus 1 to bus 2 and c is the capacity on the line.

To apply Algorithm 1, first solve the NDA-OPF (Eq. 4.16) for the two bus network. Then, to apply Step 2, we partition $\mathbb{R}^2$ into the five regions in Fig. 4.5 according to the value of the net demand forecast $\hat{\mathbf{d}}$. Each region is defined by whether the transmission link is congested or not, the direction of congestion, and whether each bus is scheduled to generate power in the nominal problem. The small-σ assumption enables inference of these facts with high probability from the solution of the NDA-OPF.

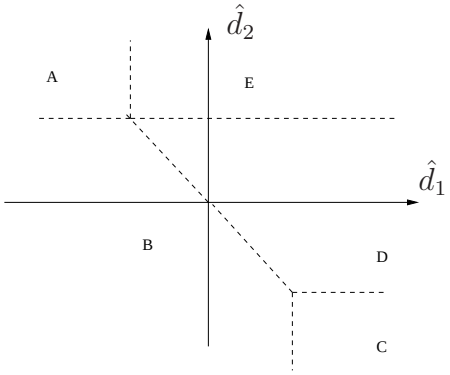


Figure 4.5: Partition of $\mathbb{R}^2$ with respect $\hat{\mathbf{d}}$ when $\alpha_1 \leq \alpha_2$. The small-σ assumption means that the actual realization of $\mathbf{d}$ is in the same region as $\hat{\mathbf{d}}$ w.h.p.

Regions A, B, C and D are reduces to the single bus case as analyzed in Section 4.2. In regions B and D, since the line capacity is not binding, power can flow from one bus to the other without congestion. In region A, bus 1 has excess power and transfer up to capacity to bus 2, and then reserve is only needed for bus 2. Region C is symmetrical to region A.

For region E in Fig. 4.5, $\hat{d}_1 > -c$ and $\hat{d}_2 > c$. Since buying at bus 1 is cheaper ($\alpha_1 < \alpha_2$), the SO should transfer up to line capacity c units of energy from bus 1 to bus 2. The NDA-OPF solution is then

$$\mathbf{g}^{\mathbf{I}} = \begin{bmatrix} \hat{d}_1 + c \\ \hat{d}_2 - c \end{bmatrix}.$$

At first glance, it seems the two buses are now decoupled and can be treated as two isolated buses since the line between them is congested. However, this viewpoint is not correct due to the *two-stage* nature of the problem and congestion being directional. In the two stage dispatch problem, the SO decides in the first stage to purchase some energy based on the

forecast and error statistics; however the actual balancing of the network occurs at the second stage. Some averaging of the errors can still occur even if the line from bus 1 to bus 2 is congested. For example, suppose that in real-time $e_1 > 0$ and $e_2 < 0$. That is, demand at bus 2 was over-predicted and demand at bus 1 was under-predicted. Due to this configuration, bus 2 needs less than c units of energy from bus 1, and the remaining energy can be utilized to satisfy the under-predicted demand in bus 1. This represents a flow from bus 2 to 1 and does not violate congestion constraints, since the line was congested from bus 1 to 2. Due to this property of opposing the congestion direction, we denominate this flow a *backflow*. For example, backflow does not arise in region A because bus 1 always has an excess of energy and does not require any energy from bus 2. Similarly for region C.

In region E, the small-σ assumption implies that $\mathbf{d} \in E$ with high probability and the line is not congested from bus 2 to bus 1 (Observation (3)). Assuming that errors e_1 and e_2 have covariance matrix

$$\Sigma_e = \sigma_e^2 \Sigma' = \sigma_e^2 \begin{bmatrix} \gamma_{11} & \rho \\ \rho & \gamma_{22} \end{bmatrix}, \tag{4.20}$$

the optimal dispatch and price of uncertainty in region E are given by:

Theorem 12. *Under the small-σ assumption, the risk limiting dispatch is given by*

$$\mathbf{g}^* = \overline{\mathbf{g}} + \sigma_e \boldsymbol{\Delta}^*,$$

where $\overline{\mathbf{g}} = [\hat{d}_1 + c \;\; \hat{d}_2 - c]^T$ and $\boldsymbol{\Delta}^*$ is the unique solution to

$$\alpha_1 = \min(\beta_1, \beta_2) \Pr(z_1 > \Delta_1, z_1 + z_2 > \Delta_1 + \Delta_2) \tag{4.21a}$$

$$\begin{aligned} \alpha_2 = {} & \beta_2 \Pr(z_2 > \Delta_2) \\ & + \min(\beta_1, \beta_2) \Pr(z_2 < \Delta_2, z_1 + z_2 > \Delta_1 + \Delta_2), \end{aligned} \tag{4.21b}$$

where $\mathbf{z} = [z_1 \; z_2]^T = \mathbf{e}/\sigma_e$. The cost of uncertainty is linear and the price of uncertainty is given by

$$\begin{aligned} p = {} & \boldsymbol{\alpha}^T \boldsymbol{\Delta}^* \\ & + \min(\beta_1, \beta_2) \{ \mathbb{E}[(z_1 + z_2 - \Delta_1^* - \Delta_2^*)^+ 1(z_2 < \Delta_2^*)] \\ & + \mathbb{E}[(z_1 - \Delta_1^*)^+ 1(z_2 > \Delta_2^*)] \} + \beta_2 \, \mathbb{E}[(z_2 - \Delta_2^*)^+]. \end{aligned} \tag{4.22}$$

Before formally proving Theorem 12, we provide an intuitive explanation of the non-linear equations in (4.21). After subtracting the nominal dispatch choice, the net demands (normalized by σ_e) are z_1 and z_2 respectively, and only backflow is allowed. The network reduces to a two bus network with a *unidirectional* link going from bus 2 to bus 1 (Fig. 4.6).

The left hand side of (4.21) can be seen as the cost of purchasing an additional unit of energy at the buses in stage 1, while the right hand side can be seen as the benefit of having that unit of energy at stage 2. Therefore (4.21) can be interpreted as balancing the cost and benefit between buying an additional of unit at stage 1. For example, one additional unit of

Figure 4.6: The perturbed network consisting of a unidirectional link and normalized demands $z_1 = e_1/\sigma_e$, $z_2 = e_2/\sigma_2$.

energy at bus 1 is useful if two event occurs: $z_1 > \Delta_1$ (bus 1 does not have enough energy) and (b) $z_1 + z_2 > \Delta_1 + \Delta_2$ (bus 2 does not have enough energy to transfer to bus 1). Since power can be transferred from bus 2 to bus 1 in the perturbed network (Fig. 4.6), the price of buying an unit of energy at real time is $\min(\beta_1, \beta_2)$ and the right hand side of (4.21a) is the expected benefit of having that unit of energy available. The price of purchasing that unit of energy at stage 1 is α_1. At optimality, equilibrium is achieved between the cost at stage 1 and the expected benefit at stage 2. Similarly, (4.21b) describes the equilibrium at bus 2

Figure 4.7 plots the ratio in the average price between a network where backflow is not taken into account and a network that allows backflow as a function of the correlation between errors e_1 and e_2. If backflow is not allowed, then the network becomes two isolated buses.

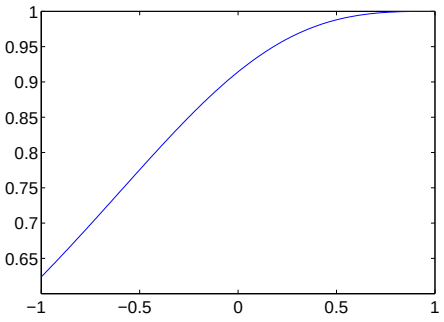


Figure 4.7: Ratio in prices between using and not using back flow for $\alpha_1 = \alpha_2 = 0.5$ and $\beta_1 = \beta_2 = 1$. Note the curve is always below one since a network with back flow can do no worse than a network without backflow.

The ratio is always less than 1 since a network with backflow can do no worse than a network without backflow. The ratio is highest when the two buses are negatively correlated since backflow averages out the uncertainties in the error. As the two buses become positively correlated, backflow becomes less useful since both errors tend to be the same sign and averaging is less useful.

Proof of Theorem 12. Any dispatch can be written as $\overline{\mathbf{g}} + \sigma_e \mathbf{\Delta}$. We first prove the optimal

$\boldsymbol{\Delta}$ is independent of σ_e. Substituting $\mathbf{g} = \overline{\mathbf{g}} + \sigma_e \boldsymbol{\Delta}$, the DA-SPF (Eqn. (4.6))becomes

$$\begin{aligned} &\text{minimize } \boldsymbol{\alpha}^T(\overline{\mathbf{g}} + \sigma_e \boldsymbol{\Delta}) + \mathbb{E}[J(\boldsymbol{\beta}, \mathbf{d} - (\overline{\mathbf{g}} + \sigma_e \boldsymbol{\Delta}))|\hat{\mathbf{d}}] && \text{(4.23a)} \\ &\text{subject to } \overline{\mathbf{g}} + \sigma_e \boldsymbol{\Delta} \geq 0. && \text{(4.23b)} \end{aligned}$$

By the small-σ assumption, the constraint in Eqn. (4.23b) is always satisfied since $\overline{\mathbf{g}} \geq 0$ from the definition of NDA-OPF. The RT-OPF (Eqn. (4.19)) becomes

$$\begin{aligned} &J(\boldsymbol{\beta}, \mathbf{d} - (\overline{\mathbf{g}} + \sigma_e \boldsymbol{\Delta})) && \text{(4.24a)} \\ =&\text{minimize } \boldsymbol{\beta}^T(\mathbf{g}^{R+1})^+ && \text{(4.24b)} \\ &\text{subject to } g_1^{R+1} + \overline{g}_1 + \sigma_e \Delta_1 - f - \hat{d}_1 - e_1 = 0 && \text{(4.24c)} \\ &\qquad g_2^{R+1} + \overline{g}_2 + \sigma_e \Delta_2 + f - \hat{d}_2 - e_2 = 0 && \text{(4.24d)} \\ &\qquad -c \leq f \leq c. && \text{(4.24e)} \end{aligned}$$

Since the nominal flow is c, let $f = c - \delta$ with δ representing the backflow. Substituting the value of $\overline{\mathbf{g}}$ into Eqn. (4.24),

$$\begin{aligned} &J(\boldsymbol{\beta}, \mathbf{d} - (\overline{\mathbf{g}} + \sigma_e \boldsymbol{\Delta})) && \text{(4.25a)} \\ =&\text{minimize } \boldsymbol{\beta}^T(\mathbf{g}^{R+1})^+ && \text{(4.25b)} \\ &\text{subject to } g_1^{R+1} + \sigma_e \Delta_1 + \delta - e_1 = 0 && \text{(4.25c)} \\ &g_2^{R+1} + \sigma_e \Delta_2 - \delta - e_2 = 0 && \text{(4.25d)} \\ &0 \leq \delta \leq 2c. && \text{(4.25e)} \end{aligned}$$

By the assumption that the line does not congest from bus 2 to bus 1, the constraint $\delta < 2c$ is always satisfied and can be dropped. Normalizing Eqn. (4.25) by σ_e gives

$$\begin{aligned} &J(\boldsymbol{\beta}, \mathbf{d} - (\overline{\mathbf{g}} + \sigma_e \boldsymbol{\Delta})) && \text{(4.26a)} \\ =&\sigma_e \text{minimize } \boldsymbol{\beta}^T(\mathbf{g}^{R+1})^+ && \text{(4.26b)} \\ &\text{subject to } g_1^{R+1} + \Delta_1 + \delta - z_1 = 0 && \text{(4.26c)} \\ &\qquad g_2^{R+1} + \Delta_2 - \delta - z_2 = 0 && \text{(4.26d)} \\ &\qquad \delta \geq 0, && \text{(4.26e)} \end{aligned}$$

where the optimization variables $\mathbf{g}^{R+1}$ and δ have been normalized by σ_e and $z_i := e_i/\sigma_e$. Let $\tilde{J} = J/\sigma_e$, and note that $\tilde{J}$ only depdent of $\boldsymbol{\beta}$ and $\boldsymbol{\Delta}$. Combining (4.23) and (4.26), $\boldsymbol{\Delta}$ solves the unconstrained optimization problem

$$\min_{\boldsymbol{\Delta}} \ \alpha^T \boldsymbol{\Delta} + E[\tilde{J}(\boldsymbol{\beta}, \boldsymbol{\Delta})]. \tag{4.27}$$

To solve this optimization problem, we need the gradient of $E[\tilde{J}(\boldsymbol{\beta}, \boldsymbol{\Delta})]$ with respect to $\boldsymbol{\Delta}$. The optimization problem can be analytically solved to yield

$$\begin{aligned}\tilde{J}(\boldsymbol{\beta}, \boldsymbol{\Delta}) &= \begin{cases} \min(\beta_1 + \beta_2)(z_1 + z_2 - \Delta_1 - \Delta_2) \\ \quad \text{if } z_1 + z_2 > \Delta_1 + \Delta_2, z_2 < \Delta_2 \\ \min(\beta_1 + \beta_2)(z_1 - \Delta_1) + \beta_2(z_2 - \Delta_2) \\ \quad \text{if } z_1 > \Delta_1, z_2 > \Delta_2 \\ \beta_2(z_2 - \Delta_2) \\ \quad \text{if } z_1 < \Delta_1, z_2 > \Delta_2 \\ 0 \text{ otherwise} \end{cases} \\ &= \min(\beta_1, \beta_2)[(z_1 + z_2 - \Delta_1 - \Delta_2)^+ 1(z_2 < \Delta_2) \\ &+ (z_1 - \Delta_1)^+ 1(z_2 > \Delta_2)] + \beta_2(z_2 - \Delta_2)^+.\end{aligned}$$

Using the linearity of expectation and taking derivatives with respect to $\boldsymbol{\Delta}$ in $\alpha^T \boldsymbol{\Delta} + E[\tilde{J}(\boldsymbol{\beta}, \boldsymbol{\Delta})]$ gives (4.21).

Next we prove the price of uncertainty is given by Eqn. (4.22). The value of full knowledge optimization problem is $\mathbb{E}[J(\boldsymbol{\alpha}, \mathbf{d})]$. The error is zero mean and by the small-σ assumption, $\mathbb{E}[J(\boldsymbol{\alpha}, \mathbf{d})] = \boldsymbol{\alpha}^T \overline{\mathbf{g}}$ where $\overline{g}$ is the nominal solution. The cost of uncertainty is

$$\begin{aligned}u &= \boldsymbol{\alpha}^T(\overline{g} + \sigma_e \boldsymbol{\Delta}) + \mathbb{E}[J(\boldsymbol{\beta}, \mathbf{d} - (\overline{\mathbf{g}} + \sigma_e \boldsymbol{\Delta}))] - \mathbb{E}[J(\boldsymbol{\alpha}, \mathbf{d})] \\ &= \sigma_e(\boldsymbol{\alpha}^T \boldsymbol{\Delta} + \mathbb{E}[\tilde{J}(\boldsymbol{\beta}, \boldsymbol{\Delta})] \\ &= \sigma_e p.\end{aligned}$$

□

4.3.2 N-bus Network with a Single Congested Line

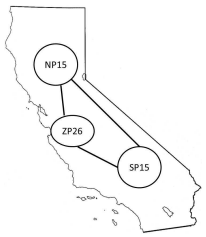


Figure 4.8: A zonal map of the California transmission network under CAISO control. The subnetwork within a zone are uncongested under normal operation. The tie lines to other WECC areas are not shown.

Most networks consists of a large number of buses and lines, but under normal operating conditions, only very few lines are congested. For example, the California transmission

network can be thought as divided into three zones connected by major transmission lines shown in Figure 4.8 and the flows within a zone are unrestricted [70]. The zonal grouping in CAISO was designed utilizing the idea of collapsing together buses connected by uncongested transmission lines in a deterministic OPF. We formalize and extend this intuitive concept for ND-RLD by showing that a general network with a single congested link reduces to a two bus problem under mild to moderate uncertainty. More concretely, assume the line from bus 1 to bus 2 is congested, then

Theorem 13. *Under the small-σ assumption, suppose that $\overline{f}_{12} = c_{12}$ is the only congestion in the network, then the following holds:*

1. *There are at most two nodes with positive generation. That is, $\overline{g}_i > 0$ for at most two i. Furthermore, if $\overline{g}_1 \leq 0$, then only one other bus has positive generation.*

2. *The optimal control takes the form*

$$g^* = (\overline{\mathbf{g}} + \mathbf{\Delta})^+,$$

   where $\Delta_i \neq 0$ only if $\overline{g}_i > 0$.

3. *If $\beta_i = \beta_k = \beta$ for all i, k, then the network can be thought as a congested two node network with congestion from bus $1'$ to $2'$ with correlated errors. Let $k \neq 1$ be the bus with positive generation. Then the first stage costs are $\alpha'_1 = \alpha_1$ and $\alpha'_2 = (\frac{\alpha_k}{\gamma_k} - \gamma_k \alpha_1)$ and the errors are given by*

$$e'_1 = e_1 + \sum_{i=3}^{n} \gamma_i e_i \qquad e'_2 = e_2 + \sum_{i=3}^{n} (1 - \gamma_i) e_i, \tag{4.28}$$

   where $\gamma_i \in [0, 1]$ are determined by the topology of the network.

Point 1) in Theorem 13 seems strange since it is highly unlikely that only two generators would be generating in a power network. This result is comes from the assumption that the prices are linear in the power generated, which is used here to simplify the presentation. In practice, cost functions are piecewise linear or quadratic. If piecewise linear cost functions are used, then Theorem 13 1) is modified to stating that there are at most two generators operating at their marginal cost [64]; if quadratic (or other continuous increasing) cost functions are used, Theorem 13 is modified to stating that there are at most two different marginal costs among the generators. The details of the derivation is given in the Appendix. The overall message of Theorem 13 remains unchanged in each case: in a network with two congested link, the risk limiting dispatch can be calculated by considering a two-bus network obtained from the original n-bus network.

The proof of this theorem is somewhat technical and can be thought as a special case of the general theorem in the next section. The theorem states that the network can be collapsed into a single bus or a two bus network, utilizing an appropriate averaging of the

net demands. To understand how to calculate the bus averaging weights γ_i, it is convenient to simplify (4.6) (with cost $\boldsymbol{\beta}$) by considering *fundamental flows* [71]. By theorems in graph theory, Eq. (4.5c) can be eliminated by choosing a spanning tree in the network, and identifying the branch flows on the tree as fundamental flows $\tilde{\mathbf{f}}$. All flows $\mathbf{f}$ are determined by these $n-1$ fundamental flows denoted $\tilde{\mathbf{f}} \in \mathbb{R}^{n-1}$. Let $\mathbf{R}$ be the mapping from $\tilde{\mathbf{f}}$ to $\mathbf{f}$ ($\mathbf{f} = \mathbf{R}\tilde{\mathbf{f}}$). Let $\mathbf{A} = \boldsymbol{\nabla}^T\mathbf{R}$. The constraint (4.5c) can be eliminated and (4.5) reduces to:

$$J^*(\boldsymbol{\beta}, \mathbf{x}) = \min \boldsymbol{\beta}^T(\mathbf{g}^R)^+ \tag{4.29a}$$

$$\text{subject to } \mathbf{g}^R - \mathbf{x} - \mathbf{A}\tilde{\mathbf{f}} = 0 \tag{4.29b}$$

$$|\mathbf{R}\mathbf{f}| \leq \mathbf{c}. \tag{4.29c}$$

Let $\mathbf{a}_i^T$ be the ith row of $\mathbf{A}$ for $i = 1, \ldots, n$. For each node $i = 3, \ldots, n$ in the network, let $\tilde{\mathbf{f}}^{(i)}$ be set of fundamental flows that solve the following set of equations

$$f_1^{(i)} = 0 \tag{4.30a}$$

$$\mathbf{a}_i^T\tilde{\mathbf{f}}^{(i)} = -1 \tag{4.30b}$$

$$\mathbf{a}_k^T\tilde{\mathbf{f}}^{(i)} = 0, \ k \neq i, k \geq 3. \tag{4.30c}$$

In matrix form, $\tilde{\mathbf{f}}^{(i)}$ solves

$$\begin{bmatrix} 1\ 0\ 0\ \cdots\ 0 \\ \mathbf{A}_2 \end{bmatrix} \tilde{\mathbf{f}}^{(i)} = \tilde{\mathbf{A}}\tilde{\mathbf{f}}^{(i)} = -\mathbf{e}_{i-1},$$

where $\mathbf{A}_2$ is the $(n-2) \times (n-1)$ matrix obtained by removing the first two rows of $\mathbf{A}$. Inverting gives $\tilde{\mathbf{f}}^{(i)} = -\tilde{\mathbf{A}}^{-1}\mathbf{e}_{i-1}$ and

$$\gamma_i = \mathbf{a}_1^T\tilde{\mathbf{f}}^{(i)}. \tag{4.31}$$

Next we apply Theorem 13 to a three bus single cycle network with equal admittance on each line. Let the prediction $\hat{\mathbf{d}}$ be such that the line from bus 1 to bus 2 is congested. That is, $\overline{f}_{12} = c_{12}$ in the nominal problem. There are four possible congestion patterns[1] as listed in Figure 4.9. Bus i is labeled by the sign of $\overline{g}_i$. Figure 4.10 shows the equivalent two bus

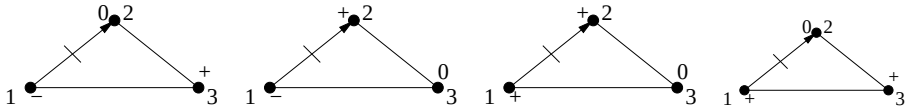


Figure 4.9: Possible sign patterns of $\overline{g}$ when a single line is congested.

networks for each of the networks in Fig. 4.9 after applying Theorem 13. The networks in Fig. 4.10 are labeled by the first stage costs, the sign patterns and the forecasted errors at each of the nodes. Let $\boldsymbol{\Delta}'$ be the solution to the two bus networks in Fig. 4.10. Then the controls $\boldsymbol{\Delta}$ for the original problem are given in each of the networks in Fig. 4.10.

[1]Other patterns are possible, but occur for a set of $\hat{d}$ that is of measure zero

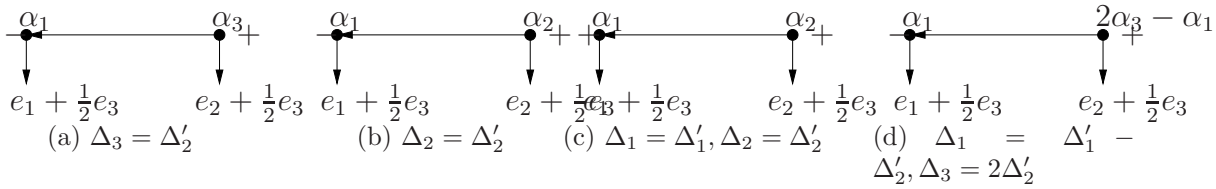


Figure 4.10: The equivalent perturbed networks for the networks in Fig. 4.9 respectively. The left bus is $1'$ and the right bus is $2'$. The back flow is only allowed form $2'$ to $1'$.

4.4 Network with Multiple Congested Links

The result from the last section carries over in an obvious way: a network with K nominally congested links can be reduced to a network with $K+1$ buses. Formally, Theorem 13 generalize to

Theorem 14. *Suppose the small-σ assumption hold. Then the two step procedure in Section 4.3 is optimal. Furthermore, let K be the number of congested links from the nominal problem. The the perturbed problem in* (4.18) *reduces to an equivalent problem on at most $K+1$ buses with at most K congested lines. The perturbed network is constructed by Algorithm 1.*

Algorithm 1: Construction of the perturbed network

1. Solve the nominal problem $J(\boldsymbol{\alpha}, \hat{\mathbf{d}})$ to get $\overline{g}$ and $\overline{\mathbf{f}}$.

2. Renumber the flows such that the K congested flows are numbered 1 to K. Let the matrix $\mathbf{A}$ be defined with respect to this numbering of the flows (2.15).

3. Let $\mathbf{A}_1$ be the matrix formed from $\mathbf{A}^T$ be keeping the first K rows and $\mathbf{A}_2$ be the matrix formed from $\mathbf{A}^T$ by removing the first K rows. Therefore $\mathbf{A}_1$ has size $K \times n$ and $\mathbf{A}_2$ has size $(n-1-K) \times n$.

4. Consider the equation
$$\mathbf{A}_2 \boldsymbol{\lambda} = \mathbf{0}. \tag{4.32}$$
From [59], all rows of $\mathbf{A}$ are linear independent, $\mathbf{A}_2$ has rank $n-K-1$ and null space of dimension of $K+1$. Therefore each λ as a linear combination of $K+1$ independent λ's.

5. The $K+1$ independent λ's include all buses with a $+$ labelling. If the number of $+$ labelled buses is less than $K+1$, the pick other buses to complete this set. Renumber the buses such that these independent buses at numbered from 1 to $K+1$. Rearrange the columns $\mathbf{A}$, $\mathbf{A}_1$ and $\mathbf{A}_2$ accordingly.

6. Find $\mathbf{B}$ such that $[\lambda_{K+2}, \ldots, \lambda_n]^T = \mathbf{B}[\lambda_1 \ldots \lambda_{K+1}]^T$. Define $\tilde{\mathbf{e}}$ and $\tilde{\mathbf{A}}$ as

$$\tilde{\mathbf{e}} = \begin{bmatrix} \mathbf{I} & \mathbf{B}^T \end{bmatrix} \mathbf{e} \qquad \tilde{\mathbf{A}} = \mathbf{A}_1 \begin{bmatrix} \mathbf{I} \\ \mathbf{B} \end{bmatrix} \tag{4.33}$$

where $\mathbf{I}$ is the identity matrix of size $K+1$, $\tilde{\mathbf{e}}$ is a vector of length $K+1$, and $\tilde{\mathbf{A}}$ is of size $K \times K+1$. Since the all one's vector spans the null space of $\mathbf{A}^T$, the rows of $\mathbf{B}$ sum up to 1.

7. The perturbed network has $K+1$ buses, with the flow to bus mapping determined by $\tilde{\mathbf{A}}^T$ and the prediction error at bus i is $\tilde{e}_i$. The topology of the reduced network can be determined by the method in Chapter 12 of [72], but is not needed in solving the perturbed problem.

8. The first and second stage prices are given by

$$\tilde{\boldsymbol{\alpha}} = \begin{bmatrix} \mathbf{I} & \mathbf{B}^T \end{bmatrix} \boldsymbol{\alpha}' \qquad \tilde{\boldsymbol{\beta}} = \begin{bmatrix} \beta'_1 & \beta'_2 & \ldots & \beta'_{K+1} \end{bmatrix}^T. \tag{4.34}$$

The perturbed two stage problem becomes

$$\min_{\tilde{\boldsymbol{\Delta}}} \ \tilde{\boldsymbol{\alpha}}^T \tilde{\boldsymbol{\Delta}} + \mathbb{E}[\tilde{J}(\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\Delta}})].$$

This problem has dimension $K+1$. Let $\tilde{\boldsymbol{\Delta}}$ the risk limiting dispatch for the perturbed problem. The original $\boldsymbol{\Delta}$ is obtained by solving

$$\min_{\boldsymbol{\Delta}} \ \boldsymbol{\alpha}'^T \boldsymbol{\Delta} \tag{4.35a}$$

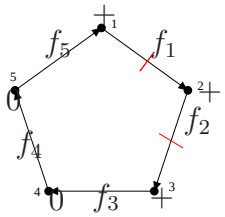
$$\text{s.t.} \ \begin{bmatrix} \mathbf{I} & \mathbf{B}^T \end{bmatrix} \boldsymbol{\Delta} = \tilde{\boldsymbol{\Delta}} \tag{4.35b}$$

$$\Delta_i \geq 0 \text{ if } \overline{g}_i = 0. \tag{4.35c}$$

The proof of Theorem 14 mainly involves showing each step in Algorithm 1 is possible and is given in the appendix. For the rest of this section, we illustrate the usefulness of Theorem 14 with examples.

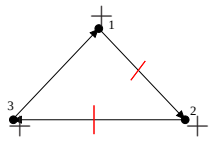
4.4.1 Example 1

Consider the 5-cycle network in Fig. 4.11. For some $\hat{\mathbf{d}}$ and α, the sign pattern of the buses after solving the nominal problem is given in Fig. 4.11. In this case renumbering of buses or the flows are not needed. Following Algorithm 1, we compute $\tilde{\mathbf{A}}$ and $\mathbf{B}$ and they are given in Fig. 4.11. The last step of Algorithm 1 gives the perturbed network is a triangle, and the direction of flows consistent with $\tilde{\mathbf{A}}$ is given in Fig. 4.12.



$$\boldsymbol{B} = \begin{bmatrix} 1/3 & 0 & 2/3 \\ 2/3 & 0 & 1/3 \end{bmatrix}$$
$$\tilde{\mathbf{A}} = \begin{bmatrix} 4/3 & -1 & -1/3 \\ 1/3 & 1 & -4/3 \end{bmatrix}$$

Figure 4.11: The resulting flows and signs of buses of a 5-cycle network after solving a nominal problem. The congested edges are label by red crosses and the directions are given by arrows (f_1 and f_2 are congested). The matrices are computed according to Algorithm 1.



$$\tilde{e}_1 = e_1 + 1/3e_4 + 2/3e_5$$
$$\tilde{e}_2 = e_2$$
$$\tilde{e}_3 = e_3 + 2/3e_4 + 1/3e_5$$
$$\tilde{\mathbf{A}} = \begin{bmatrix} 4/3 & -1 & -1/3 \\ 1/3 & 1 & -4/3 \end{bmatrix}$$

Figure 4.12: The perturbed network. The error at each bus is a linear combination of the original errors. The fundamental flows in the network is the flow from bus 1 to bus 2 and bus 2 to bus 3. The congestion is denoted by red lines.

For this perturbed network, the second stage optimization problem becomes

$$\tilde{J}(\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\Delta}}) = \min_{\mathbf{y}} \tilde{\boldsymbol{\beta}}^T \mathbf{y} \tag{4.36a}$$
$$\text{s.t. } \mathbf{y} - \tilde{\mathbf{A}}^T \mathbf{f} - (\tilde{\mathbf{e}} - \tilde{\boldsymbol{\Delta}}) \geq \mathbf{0} \tag{4.36b}$$
$$\mathbf{f} \leq \mathbf{0} \tag{4.36c}$$
$$\mathbf{y} \geq 0, \tag{4.36d}$$

where $\mathbf{f}$ is the two fundamental flows in Fig. 4.12. By standard duality theory, the dual is

$$\tilde{J}(\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\Delta}}) = \max_{\boldsymbol{\lambda}} \boldsymbol{\lambda}^T (\mathbf{e} - \tilde{\boldsymbol{\Delta}}) \tag{4.37a}$$
$$\text{s.t. } \mathbf{0} \leq \boldsymbol{\lambda} \leq \tilde{\boldsymbol{\beta}} \tag{4.37b}$$
$$\tilde{\mathbf{A}}\boldsymbol{\lambda} \leq \mathbf{0}, \tag{4.37c}$$

where $\boldsymbol{\lambda}$ are the Lagrange multipliers of the energy balance constraint (4.36b). The overall problem is

$$\min_{\tilde{\Delta}_1, \tilde{\Delta}_2, \tilde{\Delta}_3} \sum_{i=1}^{3} \tilde{\alpha}_i \tilde{\Delta}_i + \mathbb{E}[J(\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\Delta}})]. \tag{4.38}$$

To solve for the first stage control $\tilde{\boldsymbol{\Delta}}$, we can set up a system of equilibrium equations. The equations are found by considering trade-off between procuring more one unit of energy

in the first stage and the potential benefit of that unit of energy in the second stage. The cost of procuring one more unit of energy at bus i is α_i and the expected benefit of that unit of energy is given by $\mathbb{E}[\lambda_i^*(\tilde{\mathbf{e}} - \tilde{\mathbf{\Delta}})]$ [73]. To obtain the optimal $\tilde{\mathbf{\Delta}}$, we solve the system of equilibrium equations:

$$\tilde{\alpha}_i = \mathbb{E}[\lambda_i^*(\tilde{\mathbf{e}} - \tilde{\mathbf{\Delta}})], \ \forall i. \tag{4.39}$$

Since $\tilde{J}$ is a linear program, the optimal occurs at a vertex of the polyhedron defined by (4.37b) and (4.37c). The vertices can be listed [74], and for each $(\tilde{\mathbf{e}} - \tilde{\mathbf{\Delta}})$, the corresponding optimal vertex can be found be simple comparisons. For the network in Fig. 4.12, there are 6 possible vertices, and due to space constraints, we do not list them here. By comparison of the vertices,

$$\lambda_1^* = \begin{cases} \beta & \text{if } \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1/4 \\ 3/4 & 1 & 0 \end{bmatrix} (\tilde{\mathbf{e}} - \tilde{\mathbf{\Delta}}) \geq 0 \\ \beta/4 & \text{if } \begin{bmatrix} 1/4 & 0 & 1 \\ 1/4 & 0 & 0 \\ 1/4 & -1 & 0 \\ -3/4 & -1 & 0 \end{bmatrix} (\tilde{\mathbf{e}} - \tilde{\mathbf{\Delta}}) \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} \beta & \text{if } \mathbf{R}_1(\tilde{\mathbf{e}} - \tilde{\mathbf{\Delta}}) \geq 0 \\ \beta/4 & \text{if } \mathbf{R}_2(\tilde{\mathbf{e}} - \tilde{\mathbf{\Delta}}) \geq 0 \\ 0 & \text{otherwise} \end{cases}.$$

Therefore the first equilibrium equation is given by

$$\tilde{\alpha}_1 = \beta \Pr(\mathbf{R}_1(\tilde{\mathbf{e}} - \tilde{\mathbf{\Delta}}) \geq 0) + \beta/4 \Pr(\mathbf{R}_2(\tilde{\mathbf{e}} - \tilde{\mathbf{\Delta}}) \geq 0)$$

and equations for $\tilde{\alpha}_2$ and $\tilde{\alpha}_3$ can be generated in the same way. Solving this set of equations gives $\tilde{\mathbf{\Delta}}^*$. And the optimal solution to the original problem is $(\overline{g} + \mathbf{\Delta})^+$ where Δ is found by solving (4.35).

4.4.2 Example 2

Consider the IEEE 9-bus benchmark network in Fig. 4.15 [23]. For some predicted demand and cost $\boldsymbol{\alpha}$, the congestion pattern and the labelling of the buses are given in Fig. 4.15 as well. The perturbed network is given in Fig. 4.14.

The equilibrium conditions can be written down using the same procedure in Section 4.4.1.

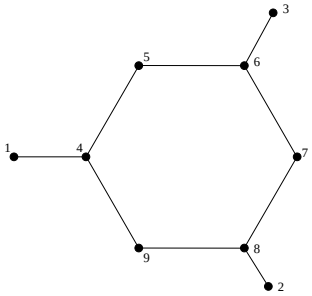
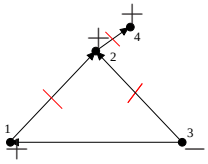


Figure 4.13: The IEEE 9-bus benchmark network with a particular sign and congestion pattern. The direction of the flows are given by the arrows, and the congested lines are denoted by red marks.



$$\tilde{e}_1 = e_1 + 1/3(e_3 + e_5) + 2/3e_6 + e_9$$
$$\tilde{e}_2 = e_2 + e_3$$
$$\tilde{e}_4 = e_7$$
$$\tilde{\mathbf{A}} = \begin{bmatrix} 4/3 & -1 & -1/3 & 0 \\ 2/3 & 1 & -5/3 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}$$

Figure 4.14: The perturbed network. The error at each bus is a linear combination of the original errors. The error at bus 3 does not matter since it is labelled $(-)$ ($\tilde{\beta} = 0$). The fundamental flows in the network is the flow from bus 1 to bus 2, bus 3 to bus 2 and bus 2 to bus 4. The congestion is denoted by red lines.

4.5 Simulation Results

This section explores various numerical examples using the IEEE 9-bus benchmark network. In particular we compare the performance of ND-RLD with utilizing the standard $3 - \sigma$ rule. We also compute the price of uncertainty numerically and compare it to the theoretical prediction.

4.5.1 Uncongested Network

Many practical networks have line capacities that are much larger than the typical power flows. For these networks, they are well approximated by a single bus network. For example, consider the IEEE 9-bus network in Figure 4.15. The nominal generation and demands from the data included with this benchmark network [23, 75] is shown in Table 4.2. Note that line flows are significantly smaller than transmission line capacities. Therefore, under moderately high penetration, the network can be thought as a network operating without

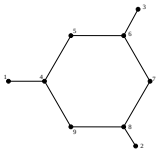


Figure 4.15: IEEE 9-bus benchmark network. Bus 1,2,3 are generators and the rest of the buses are loads.

capacity constraints.

Up to this point we have used the DC power flow model, while in reality power flow is AC. It is known that for transmission networks, due to the low R/X ratios of the transmission lines, DC and AC power flows yields similar answers. This is confirmed in our simulations where the difference in performance of using the risk limiting dispatch under DC and AC power flow models is minimal. Therefore it is sufficient to use the simpler DC flow model to obtain the dispatch.

Bus	1	2	3	4	5	6	7	8	9
DC Flow	86.6	134.4	94.1	0	-90	0	-100	0	-125
AC Flow	89.8	134.3	94.2	0	-90	0	-100	0	-125

Table 4.1: All units are MW. Negative numbers are the demands at buses 5, 7, and 9. The generations needed at buses 1, 2, and 3 to meet these demands under both DC flow and AC flow are shown. Note that due to losses, the total amount of generations needed under the AC power flow model is slightly higher.

From bus	1	4	5	3	6	7	8	8	9
To bus	4	5	6	6	7	8	2	9	4
DC Flow	86.6	33.7	-56.3	94.1	37.8	-62.2	- 134.4	72.2	-52.8
AC Flow	89.8	35.2	-55.0	94.2	38.2	-61.9	-134.3	72.11	-54.3
Capacity	250	250	150	300	150	250	250	250	250

Table 4.2: All units are MW. Both DC and AC power flows on each line of the network is shown. Capacities are the long term emergency rating of the line. The network is uncongested since at most a flow takes about half of the capacities on the line.

To analyze the performance of the risk limiting dispatch derived in Section 4.2, we compare it to two other dispatches. The first one is the currently used $3-\sigma$ dispatch, and the second one is the oracle dispatch where the actual realization of the wind is known at stage 1. We assume that all the generating buses have a first stage cost[2] $\alpha = 1$ and all buses have the same second stage cost β. For simplicity, the prediction errors are generated as i.i.d. zero mean Gaussian random variables with variance σ^2. The predictions $\hat{\mathbf{d}}$ is taken to the nominal demands in Tab. 4.1.

[2]The nominal generations are determined by an OPF problem, and every generator with non-zero generation has the same marginal cost. This can be thought as α.

The risk limiting dispatch is derived by viewing the network as a single bus. For actual operation, the amount of reserves to put at each buses in the network need to be determined. Here we spread the reserves equally among the three generating buses(buses 1,2 and 3). From (4.11) and the fact that the prediction errors are independent, the single bus risk limiting dispatch is $\sum_{i=1}^{9} \hat{d}_i + \Delta$ where $\Delta = \sqrt{9}\sigma Q^{-1}(\frac{\alpha}{\beta})$. The network risk limiting dispatch is given by

$$\begin{aligned}\mathbf{g}_{\text{rld}} = \overline{g} + \mathbf{\Delta} = &\begin{bmatrix}86.6 & 134.4 & 94.1 & 0 & \dots & 0\end{bmatrix}^T \\ &+ 3\sigma Q^{-1}(\frac{\alpha}{\beta}) \begin{bmatrix}\frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & \dots & 0\end{bmatrix}^T.\end{aligned}$$

The $3-\sigma$ control purchases a reserve of 3 times the standard deviation for each bus in the network, or $3 \cdot 9 \cdot \sigma$. Again we spread out the $3-\sigma$ dispatch over the three generating nodes as

$$\begin{aligned}\mathbf{g}_{\text{rld}} = \overline{g} + \mathbf{\Delta} = &\begin{bmatrix}86.6 & 134.4 & 94.1 & 0 & \dots & 0\end{bmatrix}^T \\ &+ 9\sigma Q^{-1}(\frac{\alpha}{\beta}) \begin{bmatrix}1 & 1 & 1 & 0 & \dots & 0\end{bmatrix}^T.\end{aligned}$$

We simulate the cost for both the DC and AC power flows.

Figure 4.16 plots the total cost of the three dispatches for $\beta = 1.5\alpha$. As we can see the risk limiting dispatch performs much better than the $3-\sigma$ dispatch. There are two reasons why the $3-\sigma$ dispatch or rules like it perform badly. The first is that the $3-\sigma$ rules is too conservative since it does not take into account the actual cost of the second stage; the second reason is that the $3-\sigma$ dispatch ignores the potential benefit of averaging between the prediction errors by treating the different buses as isolated nodes. In contrast, risk limiting dispatch takes these two points into consideration. Figure 4.17 is a zoomed in version of Fig.

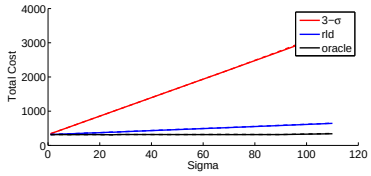


Figure 4.16: Total costs for $\beta = 1.5\alpha$ as a function of σ. The red, blue and black lines are the total cost for the $3-\sigma$, risk limiting, and the oracle dispatches respectively. The solid lines are the costs under DC flow while the dotted lines are for AC flows.

4.16 by plotting the total cost only for the risk limiting dispatch and the oracle dispatch. The cost for the oracle dispatch is constant at 315 up until $\sigma = 80$. This is expected since the predicted total demand is 315 MW, and the prediction errors are zero mean, so the errors averages out. At higher σ, the capacities in the network become binding and the cost goes up since not all errors can be averaged. The cost for the risk limiting dispatch is essentially

linear for all σ's. A lower bound for the minimum total cost is the total cost of applying the risk limiting dispatch to a network with infinite capacities, since an infinite capacity network has lower cost than a finite capacity one and the risk limiting dispatch is optimal for the former. From Figure 4.17, this lower bound is almost met. Thus the risk limiting dispatch is close to optimal and our assumption of viewing an uncongested network as a single bus network is valid. The slopes of the lines gives the price of uncertainties. As expected, the

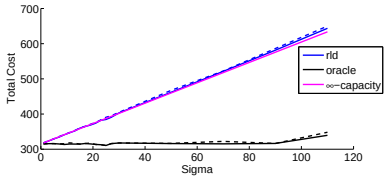


Figure 4.17: Total costs for $\beta = 1.5\alpha$ as a function of σ. The blue and black lines are the total cost for the risk limiting and the oracle dispatches respectively. The purple line is the cost of the RLD when applied to an infinite capacity network, which is a lower bound for the minimum cost of the finite capacity network. The slopes of the blue and the purple lines represent the price of uncertainties.

price of uncertainty for the oracle dispatch is 0 since the actual realization is known at the first stage. The price of uncertainty of the risk limiting dispatch closely matches that of the single bus price of uncertainty, while the $3 - \sigma$ price is much higher.

4.5.2 Congested Network

To construct a congested network, the network in Fig. 4.15 is modified by increasing the nominal load at bus 5 to 150 MW and reducing the capacity of the line connecting bus 5 and 6 to 75 MW. Then the line from bus 6 to bus 5 is congested. There are two different first stage costs α_1 and α_2 and these are given by the marginal costs of the generators. Let $\alpha = \frac{1}{2}(\alpha_1 + \alpha_2)$ and we normalize all cost by α. Figure 4.18 plots the total cost of the three dispatches for $\beta = 1.5\alpha$. Again, we see the risk limiting dispatch performs much better than the $3 - \sigma$ dispatch. Figure 4.19 is a zoomed in version of Fig. 4.18 with the total cost only for the risk limiting dispatch and the oracle dispatch. As expected, the cost of the oracle dispatch is constant over a wide range of σ's. The cost of the risk limiting dispatch is linear and very close to its lower bound. The lower bound is obtained by applying the risk limiting dispatch to a network with only one finite capacity line, namely the line congested under the nominal flows. Figure 4.19 shows that the abstraction of thinking about a network with one congested line as a two bus network is very accurate.

Figure 4.20 shows the difference in cost of using the risk limiting dispatch derived in Section 4.2 and using the risk limiting dispatch developed in this section. As expected, the later dispatch performs better since it takes into account the congestion in the network. The gap in cost should increase as the ratio $\frac{\beta}{\alpha}$ increases.

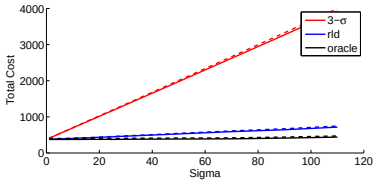


Figure 4.18: Total costs for $\beta = 1.5\alpha$ as a function of σ. The red, blue and black lines are the total cost for the $3 - \sigma$, risk limiting, and the oracle dispatches respectively. The solid lines are the costs under DC flow while the dotted lines are for AC flows. The purple line is the cost of the rld when applied to a network where only one line has finite capacity, namely the line congested under the nominal flows. This is a lower bound for the minimum cost of the finite capacity network. The slopes of the blue and the purple lines represent the price of uncertainties.

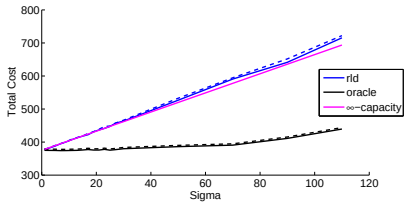


Figure 4.19: Total costs for $\beta = 1.5\alpha$ as a function of σ. The blue and black lines are the total cost for the risk limiting and the oracle dispatches respectively. The purple line is the cost of the RLD when applied to a network where only one line has finite capacity, namely the line congested under the nominal flows. This is a lower bound for the minimum cost of the finite capacity network. The slopes of the blue and the purple lines represent the price of uncertainties.

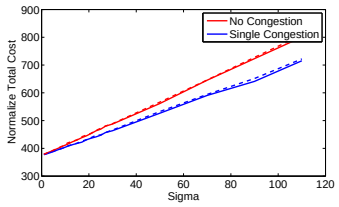


Figure 4.20: Total costs for $\beta = 1.5\alpha$ as a function of σ. The blue line is the total cost for the risk limiting dispatch developed in Section 4.3. The red line is the total cost if the risk limiting dispatch derived for the congested network in Section 4.2 is used.

Chapter 5

Conclusion

In this thesis we have considered the control and optimization of resources in the power system in the presence of renewables. We investigated the problem of optimal power flow by looking at the injection region. We explored various geometric property of the injection region, and in particular, showed that for a network of a tree topology, the Pareto-Front of the injection region is conserved under the convex hull operation. Therefore the OPF problem can be convexified for distribution networks. As an application, we derive an optimal and decentralized algorithm for the voltage distribution problem.

Next, we investigated the problem of stochastic generator dispatching in the transmission network. We formulated it as a two-stage stochastic programming problem, and showed under the small-error regime, it can be solved in an analytical fashion without sampling. Furthermore, we quantify the impact of uncertainty on the system cost with a notion called the price of uncertainty.

Bibliography

[1] National Renewable Energy Laboratory, "Solar radiation data," http://www.nrel.gov/solar/, 2011.

[2] G. Constable and B. Somerville, *A Century of Innovation: Twenty Engineering Achievement that Transformed Our Lives*. Joseph Henry Press, 2003.

[3] "Large blackouts in north america: Historical trends and policy implications," *Energy Policy*, vol. 37, no. 12, pp. 5249 – 5259, 2009.

[4] Governer of California, "Executive order s-14-08," 2008.

[5] California Engergy Commission & California Public Utilities Commission. (2013) Go solar california. [Online]. Available: http://www.gosolarcalifornia.ca.gov/

[6] Governer of California, "Executive order to help bring 1.5 million zero-emission vehicles onto californias roads," 2012.

[7] A. Bergen and V. Vittal, *Power System Analysis*. Upper Saddle River, NJ: Prentice Hall, 2000.

[8] K. Y. Lee, X. Bai, and Y. M. Park, "Optimization method for reactive power planning by using a modified simple genetic algorithm," *IEEE Transactions on Power Systems*, vol. 10, no. 4, 1995.

[9] W. C. Hahn, "Load studies on the d-c calculating table," *General Electric Review*, vol. 34, 1931.

[10] J. A. Casazza and W. S. Ku, "The coordinate use of a-c and d-c network analyzers," in *Proc. of Amer. Power Conf.*, vol. 16, Chicago, IL, 1954.

[11] B. Stott, J. Jardim, and O. Alsac, "DC power flow revisited," *IEEE Trans. on Power Sys.*, 2009.

[12] J. Carpentier, "Optimal power flows," *Int. J. Elect. Power Engergy Syst.*, no. 1, 1979.

[13] M. Huneault and F. Galiana, "A survey of the optimal power flow literature," *IEEE Transactions on Power Systems*, 1991.

[14] Z. Qiu, G. Deconinck, and R. Belmans, "A literature survey of optimal power flow problems in the electricity market context," in *Power Systems Conference and Exposition, 2009. PSCE '09. IEEE/PES*, 2009, pp. 1–6.

[15] S. Frank, I. Steponavice, and S. Rebennack, "Optimal power flow: a bibliographic survey," *Energy Systems*, 2012.

[16] S. Cvijic and M. Ilic, "Concepts of DYMONDS simulator," in *Innovative Smart Grid Technologies*, 2011.

[17] A. Trias, "The holomorphic embedding load flow method," in *IEEE Power and Energy Society General Meeting*, 2011.

[18] A. Santos, S. Deckmann, and S. Soares, "A dual lagrangian approach for optimal power flow," *IEEE Transaction on Power Systems*, vol. 3, no. 3, 1988.

[19] R. A. Jabr, A. H. Coonick, and B. J. Cory, "A primal-dual interior point method for optimal power flow dispatching," *IEEE Transaction on Power Systems*, vol. 17, no. 3, 2002.

[20] R. A. Jabr, "Optimal power flow using and extended conic quadratic formulation," *IEEE Transaction on Power Systems*, vol. 23, no. 3, 2008.

[21] X. Bai, H. Wei, K. Fujisawa, and Y. Wang, "Semidefinite programming for optimal power flow problems," *Electrical Power and Energy Systems*, 2008.

[22] J. Lavaei and S. Low, "Zero duality gap in optimal power flow problem," *Power Systems, IEEE Transactions on*, vol. 27, no. 1, pp. 92–107, 2012.

[23] University of Washington, "Power systems test case archive," http://www.ee.washington.edu/research/pstca.

[24] B. Lesieutre, D. Molzahn, A. Borden, and C. L. DeMarco, "Examining the limits of the application of semidefinite programming to power flow problems," in *Allerton 2011*, 2011.

[25] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, 2004.

[26] P. Carvalho, P. Correia, and L. Ferreira, "Distributed reactive power generation control for voltage rise mitigation in distribution networks," *IEEE Trans. Power Sys.*, vol. 23, no. 2, pp. 766–772, May 2008.

[27] A. Keane, L. Ochoa, E. Vittal, C. Dent, and G. Harrison, "Enhanced utilization of voltage control resources with distributed generation," *IEEE Trans. Power Sys.*, vol. 26, no. 1, pp. 252–260, Feb. 2011.

[28] C. Guille and G. Gross, "A conceptual framework for the vehicle-to-grid (v2g) implementation," *Energy Policy*, vol. 37, no. 11, pp. 4379–4390, Nov. 2009.

[29] G. Joos, B. Ooi, D. McGillis, F. Galiana, and R. Marceau, "The potential of distributed generation to provide ancillary services," in *Proc. of IEEE Power Engineering Society Summer Meeting*, Seattle, WA, 2000.

[30] D. Logue and P. Krein, "Utility distributed reactive power control using correlation techniques," in *Proc. of IEEE Applied Power Electronics Conference*, Anaheim, CA, March 2001.

[31] Petra Solar. (2009) SunWave Pole-Mount Solutions. South Plainfield, NJ. [Online]. Available: http://www.petrasolar.com/

[32] SolarBridge Technologies. (2009) Pantheon Microinverter. South Plainfield, NJ. [Online]. Available: http://www.petrasolar.com/

[33] M. Baran and I. El-Markabi, "A multiagent-based dispatching scheme for distributed generators for voltage support on distribution feeders," *IEEE Trans. Power Syst.*, vol. 22, no. 1, pp. 52–59, Feb. 2007.

[34] K. Turitsyn, P. Sulc, S. Backhaus, and M. Chertkov, "Distributed control of reactive power flow in a radial distribution circuit with high photovoltaic penetration," in *Proc. of IEEE Power and Energy Society General Meeting, 2010*, Minneapolis, MN, July 2010, pp. 1–6.

[35] D. Villacci, G. Bontempi, and A. Vaccaro, "An adaptive local learning-based methodology for voltage regulation in distribution networks with dispersed generation," *IEEE Trans. Power Sys.*, vol. 21, no. 3, pp. 1131–1140, Aug. 2006.

[36] A. A. Aquino-Lugo, R. Klump, and T. J. Overbye, "A control framework for the smart grid for voltage support using agent-based technologies," *IEEE Trans. on Smart Grid*, vol. 2, no. 1, pp. 173 –180, March 2011.

[37] N. P. Padhy, "Unit commitmenta bibliographical survey," *IEEE Transaction on Power Systems*, vol. 19, no. 2, 2004.

[38] S. Salam, "Unit commitment solution methods," in *PROCEEDINGS OF WORLD ACADEMY OF SCIENCE, ENGINEERING AND TECHNOLOGY*, vol. 26, 2007.

[39] S. Takriti, J. Birge, and E. Long, "A stochastic model for the unit commitment problem," *Power Systems, IEEE Transactions on*, vol. 11, no. 3, pp. 1497–1508, 1996.

[40] A. Papavasiliou, S. Oren, and R. P. O'Neill, "Reserve requirements for wind power integration: A scenario-based stochastic programming framework," vol. 26, pp. 2197–2206, Nov. 2011.

[41] A. Tuohy, P. Meibom, E. Denny, and M. O'Malley, "Unit commitment for systems with significant wind integration," *IEEE Transactions on Power Systems*, vol. 24, pp. 592–601, May 2009.

[42] P. A. Ruiz, C. R. Philbrick, E. Zak, K. W. Cheung, and P. W. Sauer, "Uncertainty management in the unit commitment problem," *IEEE Transactions on Power Systems*, vol. 24, pp. 642–651, May 2009.

[43] N. Growe-Kuska, H. Heitsch, and W. Romisch, "Scenario reduction and scenario tree construction for power management problems," in *Proceedings of the IEEE Bologna Power Tech Conference Proceedings*, vol. 3, Jun. 2003.

[44] Q. P. Zheng, J. Wang, P. M. Pardalos, and Y. Guan, "A decomposition approach to the two-stage stochastic unit commitment problem," *Ann. Oper. Res.*, 2012.

[45] D. Bertsimas, E. Litvinov, X. A. Sun, J. Zhao, and T. Zheng, "Adaptive robust optimization for the security constrained unit commitment problem," *Submitted to IEEE Transactions on Power Systems*, 2011.

[46] L. Lu, J. Tu, C.-K. Chau, M. Chen, and X. Lin, "Online energy generation scheduling for microgrids with intermittent energy sources and co-generation," *ArXiv:abs/1211.4473*, 2012. [Online]. Available: http://arxiv.org/abs/1211.4473

[47] W. H. Kersting, *Distribution system modeling and analysis*. CRC Press, 2006.

[48] J. Jarjis and F. Galiana, "Quantitative analysis of steady state stability in power systems," *IEEE Trans. Power App. Syst.*, 1981.

[49] W. Hogan, "Contract networks for electric power transmission: Technical reference," http://ksghome.harvard.edu/˜whogan/acnetref.pdf, 1992.

[50] Distribution Test Feeder Working Group, "Distribution test feeders," http://ewh.ieee.org/soc/pes/dsacom/testfeeders/index.html, 2010.

[51] D. Fudenberg and J. Tirole, *Game Theory*. MIT Press, 1991.

[52] B. Zhang and D. Tse, "Geometry of injection regions of power networks," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 788–797, 2013.

[53] S. Sojoudi and J. Lavaei, "Network topologies guaranteeing zero duality gap for optimal power flow problem," in *Proc. of IEEE PES General Meetings*, 2012.

[54] S. Bose, D. F. Gayme, S. Low, and M. K. Chandy, "Optimal power flow over tree networks," in *In Proc. of the Forth-Ninth Annual Allerton Conference*, Monticello, IL, 2011.

[55] R. Baldick, *Applied Optimization: Formulation and Algorithms for Engineering Systems*. Cambridge, 2006.

[56] J. Lavaei and S. Low, "Zero duality gap in optimal power flow," *IEEE Trans. on Power Syst.*, vol. 27, no. 1, Feb. 2012.

[57] I. Hiskens and R. Davy, "Exploring the power flow solution space boundary," *IEEE Transactions on Power Systems*, vol. 16, no. 3, pp. 389–395, 2001.

[58] D. A. Spielman, "Lecture notes on spectral graph theory," http://www.cs.yale.edu/homes/spielman/561/lect02-09.pdf, 2009.

[59] R. Diestel, *Graph Theory*. Springer-Verlag, 2010.

[60] K. Zou, A. P. Agalgaonkar, K. M. Muttaqi, and S. Perera, "Distribution system planning with incorporating dg reactive capability and system uncertainties," *IEEE Trans. on Sustainable Engergy*, vol. 3, no. 1, 2012.

[61] E. A. J. Brea, "Simple photovaltic solar cell dynamic sliding mode controlled maimum power point tracker for battery charging applications," in *Proceedings of the Twenty-Fifth Annual IEEE Applied Power Electronics Conference and Exposition (APEC)*, Feb. 2010, pp. 666–671.

[62] R. E. Tarjan and M. Yannakakis, *Simple linear-time algorithms to test chordality of graphs, test acyclicity of hypergraphs, and selectively reduce acyclic hypergraphs*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, July 1984, vol. 13.

[63] H. Terelius, U. Topcu, and R. M. Murray, "Decentralized multi-agent optimization via dual decompostion," in *18th World Congress of the International Federation of Automatic Control*, 2010.

[64] D. S. Kirschen and G. Strbac, *Fundamentals of power system economics*, 2004.

[65] R. Rajagopal, J. Bialek, C. Dent, R. Entriken, F. F. Wu, and P. Varaiya, "Risk limiting dispatch: Empirical study," in *12th International Conference on Probabilistic Methods Applied to Power Systems*, 2012.

[66] R. Rajagopal, E. Bitar, F. F. Wu, and P. Varaiya, "Risk-Limiting Dispatch for Integrating Renewable Power," *International Journal of Electrical Power and Energy Systems, to appear*, 2012.

[67] ——, "Risk Limiting Dispatch of Wind Power," in *Proceedings of the American Control Conference (ACC)*, 2012.

[68] A. B. Kurzhanski and P. Varaiya, "Dynamic optimization for reachability problems," *J. Optim. Theory Appl.*, vol. 108, no. 2, pp. 227–251, feb 2001.

[69] H. Theil, "A note on certainty equivalence in dynamic planning," *Econometrica*, 1957.

[70] Department of Market Monitoring, "Annual report on market issues and performance," California ISO, Tech. Rep., 2007.

[71] J. A. Bondy and U. S. R. Murty, *Graph Theory*. Springer, 2008.

[72] ——, *Graph Theory with Applications*. Elsevier Science, 1976.

[73] D. Bertsimas and J. Tsitsikilis, *Introduction to Linear Optimization*. Athena Scientific, 1997.

[74] D. Avis and K. Fukuda, "A pivoting algorithm for convex hulls and vertex enumeration of arrangements and polyhedra," *Discr. Comput. Geom.*, 1992.

[75] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER's extensible optimal power flow architecture," in *IEEE Power and Energy Society General Meeting*, 2009.

[76] R. A. Horn and C. A. Johnson, *Matrix Analysis*. Cambridge, 1985.

[77] H. L. Bodlaender, "A linear time algorithm for finding tree-decompositions of small treewidth," *SIAM Journal on Computing*, 1996.

[78] S. M. Fallat and L. Hogben, "The minimum rank of symmetric matrices described by a graph: A survey," *Linear Algebra Applications*, 2007.

[79] H. van der Holst, "Graphs whose positive semi-definite matrices have nullity at most two," *Linear Algebra and its Applications*, 2003.

[80] C. R. Johnson, R. Loewy, and P. A. Smith, "The graphs for which the maximum multiplicity of an eigenvalue is two," *Linear and Multilinear Algebra*, 2009.

[81] J. Beagley, E. Radzwion, S. Rimer, R. Tomasino, J. Wolfe, and A. Zimmer, "On the minimum semidefinite rank of a graph using vertex sums, graphs with $msr(G) = |G| - 2$, and the msrs of certain graph classes," NSF-REU report from Central Michigan University, 2007.

Appendix A

Proofs of Results in Chapter 2

A.1 Proof of Theorem 1

The following basic lemma from linear algebra is useful.

Lemma 15 (Rank Nullity Theorem)**.** *Let A be a $n \times n$ real symmetric matrix. Let* $\operatorname{image}(A)$ *and* $\ker(A)$ *denote the image and kernel of A, respectively. Then* $\dim \operatorname{image}(A) + \dim \ker(A) = n$ *and* $\operatorname{image}(A) \oplus \ker(A) = \mathbb{R}^n$*, where $\oplus$ is the direct sum.*

First consider the case where the network is lossless. Then any feasible injection vector must be on the conservation of energy plane. We need to show that any point on the plane can be achieved. Since the network is lossless $\mathbf{Y} = j \operatorname{Im}(\mathbf{Y})$ where $\operatorname{Im}(\mathbf{Y})$ is a $n \times n$ real symmetric matrix and each row of $\operatorname{Im}(\mathbf{Y})$ sums to 0 by (2.1). Therefore $\operatorname{Im}(\mathbf{Y})$ is a generalized graph Laplacian matrix where the admittances can be interpreted as weights on the edges. By a standard result in graph theory, $\dim \ker(\operatorname{Im}(\mathbf{Y})) = 1$ and $\ker(\operatorname{Im}(\mathbf{Y}))$ is spanned by the all one's vector $\mathbf{1}$. By Lemma 15, $\operatorname{image}(\operatorname{Im}(\mathbf{Y}))$ is the linear subspace in $\mathbb{R}^n$ orthogonal to $\mathbf{1}$. Let $\mathbf{p}^0$ be an injection vector on the conservation of energy plane, that is $\sum_{i=1}^n P_i^0 = 0$. Since $\mathbf{1}^T \mathbf{p}^0 = 0$, there is a unique vector $\mathbf{v}^0$ such that $\mathbf{Y}\mathbf{v}^0 = \mathbf{p}^0$ and $\mathbf{1}^T \mathbf{v}^0 = 0$. Choose the voltage vector $\mathbf{v} = (-\mathbf{v}^0 + j\mathbf{1})$, then

$$
\begin{aligned}
&\operatorname{Re}(\operatorname{diag}((-\mathbf{v}^0 + j\mathbf{1})(-\mathbf{v}^0 + j\mathbf{1})^H \mathbf{Y}^H)) && \text{(A.1)}\\
&= \operatorname{Re}(\operatorname{diag}((\mathbf{v}^0 \mathbf{1}^T + \mathbf{1}(\mathbf{v}^0)^T) \operatorname{Im}(\mathbf{Y})) && \text{(A.2)}\\
&\quad + j \operatorname{diag}((\mathbf{v}^0 (\mathbf{v}^0)^T + \mathbf{1}\mathbf{1}^T) \operatorname{Im}(\mathbf{Y})))\\
&\overset{(a)}{=} \mathbf{p}^0,
\end{aligned}
$$

where (a) follows from the choice of $\mathbf{v}^0$ and $\operatorname{Im}(\mathbf{Y})$ being symmetric. This finishes the proof for a lossless network.

Next consider the case where the network is lossy. The proof proceeds in two parts, first we show that the conservation of energy boundary $\sum_{i=1}^n P_i = 0$ can be arbitrarily closely from above, and then we show the injection region is convex. Since the network is lossy,

$\mathrm{Re}(\mathbf{Y})$ is a $n \times n$ real positive semidefinite Laplacian matrix. By conservation of energy, any power injection vector achieved must satisfy $\sum_{i=1}^{n} P_i > 0$ if $\mathbf{p} \neq 0$. Let $\mathbf{p}^0$ be a vector on the conservation of energy plane. We show there is a voltage vector $\mathbf{v}$ that achieves a point arbitrarily close to $\mathbf{p}^0$. Since $\mathbf{1}^T\mathbf{p}^0 = 0$, by Lemma 15 there is a unique vector $\mathbf{v}^0$ such that $\mathrm{Re}(\mathbf{Y})\mathbf{v}^0 = \mathbf{p}^0$ and $\mathbf{1}^T\mathbf{v}^0 = 0$. Let $\mathbf{v} = (\alpha\mathbf{1} + \frac{1}{\alpha}\mathbf{v}^0)$ for some $\alpha \geq 0$ and the corresponding injection vector $\mathbf{p}$ is

$$\begin{aligned}
\mathbf{p} &= \mathrm{Re}(\mathrm{diag}(\mathbf{v}\mathbf{v}^T\mathbf{Y})) \qquad (A.3)\\
&= \mathrm{Re}(\mathrm{diag}((\alpha\mathbf{1} + \frac{1}{\alpha}\mathbf{v}^0)(\alpha\mathbf{1} + \frac{1}{\alpha}\mathbf{v}^0)^T(\mathrm{Re}(\mathbf{Y}) + j\,\mathrm{Im}(\mathbf{Y})))\\
&= \mathrm{diag}((\alpha^1\mathbf{1}\mathbf{1}^T + \mathbf{v}^0\mathbf{1}^T + \mathbf{1}(\mathbf{v}^0)^T + \frac{1}{\alpha^2}\mathbf{v}^0(\mathbf{v}^0)^T)\,\mathrm{Re}(\mathbf{Y}))\\
&\overset{(a)}{=} \mathrm{diag}(\mathbf{1}(\mathbf{v}^0)^T\,\mathrm{Re}(\mathbf{Y})) + \frac{1}{\alpha^2}\,\mathrm{diag}(\mathbf{v}^0(\mathbf{v}^0)^T\,\mathrm{Re}(\mathbf{Y}))\\
&\overset{(b)}{=} \mathrm{diag}(\mathbf{1}(\mathbf{p}^0)^T) + \frac{1}{\alpha^2}\,\mathrm{diag}(\mathbf{v}^0(\mathbf{p}^0)^T)\\
&= \mathbf{p}^0 + \frac{1}{\alpha^2}\,\mathrm{diag}(\mathbf{v}^0(\mathbf{p}^0)^T),
\end{aligned}$$

where (a) follows from $\mathbf{1} \in \ker(\mathrm{Re}(\mathbf{Y}))$ and $\mathrm{Re}(\mathbf{Y})$ is symmetrical, (b) follows from the choice of $\mathbf{v}^0$. We can increase α to make $\mathbf{p}$ arbitrarily close to $\mathbf{p}^0$. For example, if we want $||\mathbf{p} - \mathbf{p}^0||_\infty \leq \epsilon$, then choose

$$\alpha \geq \sqrt{\frac{||\mathbf{p}^0||_\infty ||\mathbf{v}^0||_\infty}{\epsilon}}.$$

The next lemma states that $\mathcal{P}$ is convex.

Lemma 16. *The injection region $\mathcal{P}$ as defined in eqn. (2.8) is a convex set.*

Before proving Lemma 16, we show Theorem 1 follows from it. The geometric intuition is illustrated in Fig. A.1 for a two bus network. The main idea is to chose a set of points sufficiently close to the conservation of energy boundary (but not on it) such that the convex cone formed by the chosen set of points with the origin includes all other points in the upper half space. Formally, for a n-bus network, define $\hat{\mathbf{p}}^{(i)} \in \mathbb{R}^n$ to be the vector with $n-1$ in the i'th position and -1 in all other positions. Each $\hat{\mathbf{p}}^{(i)}$ satisfies $\mathbf{1}^T\hat{p}^{(i)} = 0$. Therefore the point $\hat{\mathbf{p}}^{(1)} + \boldsymbol{\epsilon}$ is in $\mathcal{P}$ for some $\boldsymbol{\epsilon}$ where $\mathbf{1}^T\boldsymbol{\epsilon} > 0$. Suppose $\mathbf{p} \in \mathbb{R}^n$ is a point in the upper half space, that is, $\mathbf{1}^T\mathbf{p} > 0$. To show $\mathbf{p}$ is in $\mathcal{P}$, we first show that $\mathbf{p}$ is in the strict interior of the convex cone formed by the origin $\mathbf{0}$ together with $\{\hat{\mathbf{p}}^{(1)} + \boldsymbol{\epsilon}, \hat{\mathbf{p}}^{(2)}, \ldots, \hat{\mathbf{p}}^{(n)}\}$. By definition, we need to show the existence of positive coefficients $\lambda_1, \ldots, \lambda_n$ such that

$$\mathbf{p} = \lambda_1(\hat{\mathbf{p}}^{(1)} + \boldsymbol{\epsilon}) + \sum_{i=2}^{n} \lambda_i \hat{\mathbf{p}}^{(i)}. \qquad (A.4)$$

Due to the perturbation $\boldsymbol{\epsilon}$, the set of vectors $\{\hat{\mathbf{p}}^{(1)} + \boldsymbol{\epsilon}, \hat{\mathbf{p}}^{(2)}, \ldots, \hat{\mathbf{p}}^{(n)}\}$ is linearly independent. Therefore λ_i's exist. To show they are positive, sum both side of (A.4) to obtain

$$\mathbf{1}^T\mathbf{p} = \lambda_1\mathbf{1}^T\boldsymbol{\epsilon}.$$

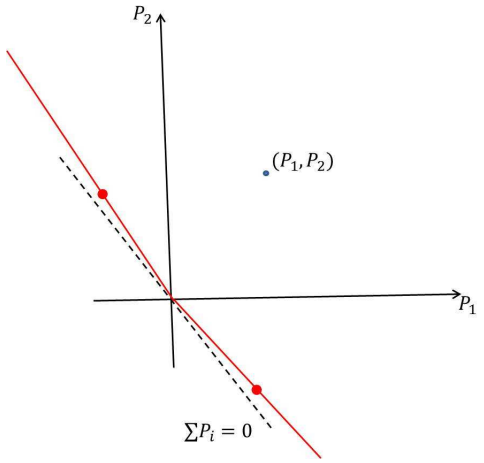


Figure A.1: Outline of proof idea for Theorem 1 for a two bus network. First pick two points close to the boundary $P_1 + P_2 = 0$ that are on opposite sides of the origin. Form the convex cone defined by the origin and the two chosen points as illustrated by the red lines. If the $\mathcal{P}$ is convex, then all points inside the cone is achievable. Since the $P_1 + P_2 = 0$ boundary can be approached arbitrarily closely, the cone can be made to include every point in the open set $\{(P_1, P_2) : P_1 + P_2 > 0\}$.

By assumptions, $\mathbf{1}^T\mathbf{p} > 0$ and $\mathbf{1}^T\boldsymbol{\epsilon} > 0$ and $\lambda_1 = \frac{\mathbf{1}^T\mathbf{p}}{\mathbf{1}^T\boldsymbol{\epsilon}} > 0$. The rest of λ_i's solves

$$\begin{bmatrix} n-1 & -1 & \cdots & -1 \\ -1 & n-1 & \cdots & -1 \\ & & \vdots & \\ -1 & -1 & \cdots & n-1 \end{bmatrix} \begin{bmatrix} \lambda_2 \\ \lambda_3 \\ \vdots \\ \lambda_n \end{bmatrix} = \begin{bmatrix} P_2 \\ P_3 \\ \vdots \\ P_n \end{bmatrix} + \begin{bmatrix} 1-\epsilon_2 \\ 1-\epsilon_3 \\ \vdots \\ 1-\epsilon_n \end{bmatrix} \lambda_1.$$

Choose an $\boldsymbol{\epsilon}$ such that $\mathbf{1}^T\boldsymbol{\epsilon}$ is sufficiently small and the right hand side of the equation is positive in every coordinate. Suppose at least one of the λ_i's are non-positive, WLOG, assume it is λ_2. The first row of the equation implies least one of $\lambda_3, \ldots, \lambda_n$ is negative . WLOG, assume it is λ_3. Furthermore, $\lambda_3 < \frac{n-1}{n-2}\lambda_2$. Repeating this procedure by going down the rows of the equation, from the second last row we get $(n-1)\lambda_n < (\lambda_2 + \cdots + \lambda_{n-1})$. But then the last row of the equation is not satisfied and therefore $\lambda_i > 0$ for all $i = 1, \ldots, n$. Since $\mathbf{p}$ is in the strict interior of the convex cone formed by $\{\hat{\mathbf{p}}^{(1)} + \boldsymbol{\epsilon}, \hat{\mathbf{p}}^{(2)}, \ldots, \hat{\mathbf{p}}^{(n)}\}$, it is in the interior of the convex cone formed by $\{\hat{\mathbf{p}}^{(1)} + \boldsymbol{\epsilon}, \hat{\mathbf{p}}^{(2)} + \boldsymbol{\epsilon}^{(2)}, \ldots, \hat{\mathbf{p}}^{(n)} + \boldsymbol{\epsilon}^{(n)}\}$ for sufficiently small $\boldsymbol{\epsilon}^{(i)}$, $i = 2, \ldots, n$. Thus $\mathbf{p}$ is in $\mathcal{P}$.

It remains to prove Lemma 16.

Proof. For a given network with n buses represented by $\mathbf{Y}$, define $\mathcal{P}_{\overline{V}}$ as

$$\mathcal{P}_{\overline{V}} = \{\mathbf{p} \in \mathbb{R}^n : \mathbf{p} = \mathrm{Re}(\mathrm{diag}(\mathbf{v}\mathbf{v}^H\mathbf{Y}^H)), ||\mathbf{v}||_2 \leq \overline{V}\}, \tag{A.5}$$

where $||\mathbf{v}||_2 = (\sum_{i=1}^n |V|_i^2)^{\frac{1}{2}}$. $\mathcal{P}_{\overline{V}}$ approaches the unconstrained injection region as $\overline{V}$ tends to infinity. $\mathcal{P}_{\overline{V}}$ cannot have holes since if $\mathbf{p} \in \mathcal{P}_{\overline{V}}$, then $\alpha\mathbf{p} \in \mathcal{P}_{\overline{V}}$ for $\alpha \in [0,1]$. Therefore to prove the convexity of $\mathcal{P}_{\overline{V}}$ it suffices to prove it has convex boundary. Consider the optimization problem

$$\begin{aligned} J = \text{minimize } & \sum_{i=1}^n c_i P_i \\ \text{subject to } & ||\mathbf{v}||_2 \leq \overline{V} \\ & \mathbf{p} = \text{Re}(\text{diag}(\mathbf{v}\mathbf{v}^H Y^H)). \end{aligned} \tag{A.6}$$

Relaxing and eliminating $\mathbf{p}$, we get

$$\begin{aligned} J_1 = \text{minimize } & \text{Tr}(\mathbf{M}\mathbf{W}) \\ \text{subject to } & \sum_{i=1} W_{ii} \leq \overline{V}^2 \\ & \mathbf{W} \succcurlyeq 0, \end{aligned} \tag{A.7}$$

By changing the costs, we are exploring the boundaries of the two regions with linear functions. We want to show that all the point on the boundary of the larger region is in fact in the smaller region.

First we show that for all $\mathbf{M}$ there is an optimal $\mathbf{W}^*$ for (A.7) which is rank 1. To solve (A.7), expand $\mathbf{W}$ in terms of its eigenvectors, so $\mathbf{W} = w_1\mathbf{w}_1\mathbf{w}_1^H + \cdots w_n\mathbf{w}_n\mathbf{w}_n^H$ where $\mathbf{w}_i$ is unit norm and $\sum_{i=1}^n w_i \leq \overline{V}^2$. Then (A.7) can be written as

$$\begin{aligned} \text{minimize } & \sum_{i=1}^n w_i \mathbf{w}_i^H M \mathbf{w}_i \\ \text{subject to } & \sum_{i=1}^n w_i \leq \overline{V}^2 \\ & \mathbf{W} = \sum_{i=1}^n (w_i \mathbf{w}_i \mathbf{w}_i^H) \succcurlyeq 0. \end{aligned} \tag{A.8}$$

By the well known result about Rayleigh quotients [76], to minimize any of the terms $\mathbf{w}_i^H M \mathbf{w}_i$, the optimal $\mathbf{w}_i^* = \mathbf{m}_1$, where $\mathbf{m}_1$ is the eigenvector corresponding to the smallest eigenvector of M. Therefore the optimal solution to (A.7) is $\mathbf{W} = \sum_{i=1}^n w_i \mathbf{m}_1 \mathbf{m}_1^H = \overline{V}^2 \mathbf{m}_1 \mathbf{m}_1^H$ and is rank 1.

If $\mathbf{m}_1$ is not unique, since eigenvector are not continuous in the entries of the matrix, we can perturb $\mathbf{Y}$ by an arbitrarily small amount to obtain a M that has a unique eigenvector corresponding to the smallest value. Note the power vector $\mathbf{p}$ is continuous in the entries of Y. From uniqueness of $\mathbf{m}_1$ and the fact there is no gap between (A.6) and (A.7), the two regions have the same boundary. Taking $\overline{V}$ to infinity finishes the proof. □

A.2 Proof of Lemma 5

To prove Lemma 5, we first show that $\mathcal{O}(\mathcal{S}) \subseteq \mathcal{O}(\mathcal{P})$ (recall that $\mathcal{S} = \text{conv}(\mathcal{P}_\theta) \cap \mathcal{P}_P$). Consider a point $\mathbf{x} \in \mathcal{O}(\mathcal{S})$, and denote its corresponding line flow from bus i to bus k with x_{ik} for every $(i,k) \in \mathcal{E}$. Due to the relation $\mathcal{P} = \mathcal{O}(\mathcal{P})$ derived in Part 2 of Theorem 4, it is enough to prove that $\mathbf{x} \in \mathcal{P}$. Since $\mathbf{x}$ is a Pareto point of the convex set $\mathcal{S}$, it is the solution of the following optimization

$$\begin{aligned} \mathbf{x} = \underset{\mathbf{p} \in \text{conv}(\mathcal{P}_\theta)}{\arg\min} & \sum_{i=1}^{n} c_i P_i \\ \text{subject to } & \underline{P}_i \leq P_i \leq \overline{P}_i \quad i = 1, 2, ..., n. \end{aligned}$$

for some positive vector $(c_1, \ldots, c_n)$. To simplify the proof, assume that all entries of this vector are strictly positive (the idea to be presented next can be adapted to tackle the case with some zero entries). By the duality theory, there exist nonnegative Lagrange multipliers $\underline{\lambda}_1, ..., \underline{\lambda}_n$ and $\overline{\lambda}_1, ..., \overline{\lambda}_n$ such that

$$\mathbf{x} = \underset{\mathbf{p} \in \text{conv}(\mathcal{P}_\theta)}{\arg\min} \sum_{i=1}^{n} (c_i + \overline{\lambda}_i - \underline{\lambda}_i) P_i - \overline{\lambda}_i \overline{P}_i + \underline{\lambda}_i \underline{P}_i$$

or equivalently

$$\mathbf{x} = \underset{\mathbf{p} \in \text{conv}(\mathcal{P}_\theta)}{\arg\min} \sum_{i=1}^{n} \left(\bar{c}_i \sum_{k \in \mathcal{V}:\ k \sim i} P_{ik} \right) \tag{A.9}$$

where $\bar{c}_i = c_i + \overline{\lambda}_i - \underline{\lambda}_i$ for every $i \in \mathcal{V}$. By complementary slackness, whenever $\bar{c}_i$ is less than or equal to zero, the multiplier $\underline{\lambda}_i$ must be strictly positive. Therefore

$$x_i = \underline{P}_i \quad \text{whenever} \quad \bar{c}_i \leq 0 \tag{A.10}$$

On the other hand, since $\text{conv}(\mathcal{P}_\theta) = \mathbf{A}\,\text{conv}(\mathcal{F}_\theta)$ and $\mathcal{F}_\theta = \prod_{(i,k) \in \mathcal{E}} \mathcal{F}_{\theta_{ik}}$, it results from (A.9) that

$$(x_{ik}, x_{ki}) = \underset{(P_{ik}, P_{ki}) \in \text{conv}(\mathcal{F}_{\theta\, ik})}{\arg\min} \bar{c}_i P_{ik} + \bar{c}_k P_{ki} \tag{A.11}$$

for every $(i,k) \in \mathcal{E}$. In order to prove $\mathbf{x} \in \mathcal{P}$, it suffices to show that $(x_{ik}, x_{ki}) \in \mathcal{F}_{\theta ik}$. Notice that if either $\bar{c}_i > 0$ or $\bar{c}_k > 0$, then it can be easily inferred from (A.11) that $(x_{ik}, x_{ki}) \in \mathcal{F}_{\theta ik}$. The challenging part of the proof is to show the validity of this relation in the case when $\bar{c}_i, \bar{c}_k \leq 0$. Consider an arbitrary vector $\mathbf{y}$ (not necessarily distinct from $\mathbf{x}$) belonging to $\mathcal{P}$. Since $(y_{ik}, y_{ki}) \in \mathcal{F}_{\theta ik}$, it is enough to prove that $(x_{ik}, x_{ki}) = (y_{ik}, y_{ki})$ whenever $\bar{c}_i, \bar{c}_k \leq 0$. This will be shown below.

Consider an edge $(i,k) \in \mathcal{E}$ such that $\bar{c}_i, \bar{c}_k \leq 0$. There exists at least one connected, induced subtree of the network including the edge (i,k) with the property that $\bar{c}_r \leq 0$ for every vertex r of this subtree. Among all such subtrees, let $\mathcal{G}$ denote the one with the

maximum number of vertices. We define two types of nodes in $\mathcal{G}$. A node $r \in \mathcal{G}$ is called a boundary node of $\mathcal{G}$ if either it is connected to some node $l \in \mathcal{V}\backslash\mathcal{G}$ or it is a leaf of the tree. We also say that a node $r \in \mathcal{V}\backslash\mathcal{G}$ is a neighbor of $\mathcal{V}$ if it is connected to some node in $\mathcal{V}$. By (A.10), if r is a node of $\mathcal{G}$, then $y_r \geq \underline{P}_r = x_r$. Without loss of generality, assume that the tree is rooted at a boundary node of $\mathcal{G}$, namely node 1.

Consider an edge (r, l) of the subtree $\mathcal{G}$ such that node l is a leaf of $\mathcal{G}$ and node r is its parent. First, we want to prove that $y_{lr} \geq x_{lr}$. To this end, consider two possibilities. If l is a leaf of the original tree, then the inequality (A.10) yields $y_{lr} = y_l \geq \underline{P}_l = x_l = x_{lr}$. As the second case, assume that l is not a leaf of the original tree. Let m denote a neighbor of $\mathcal{G}$ connected to l. By analyzing the flow region for the line (l, m) as depicted in Figure A.2a, it follows from (A.11) and the inequalities $c_l \leq 0, c_m > 0$ that (x_{lm}, x_{ml}) is at the lower right

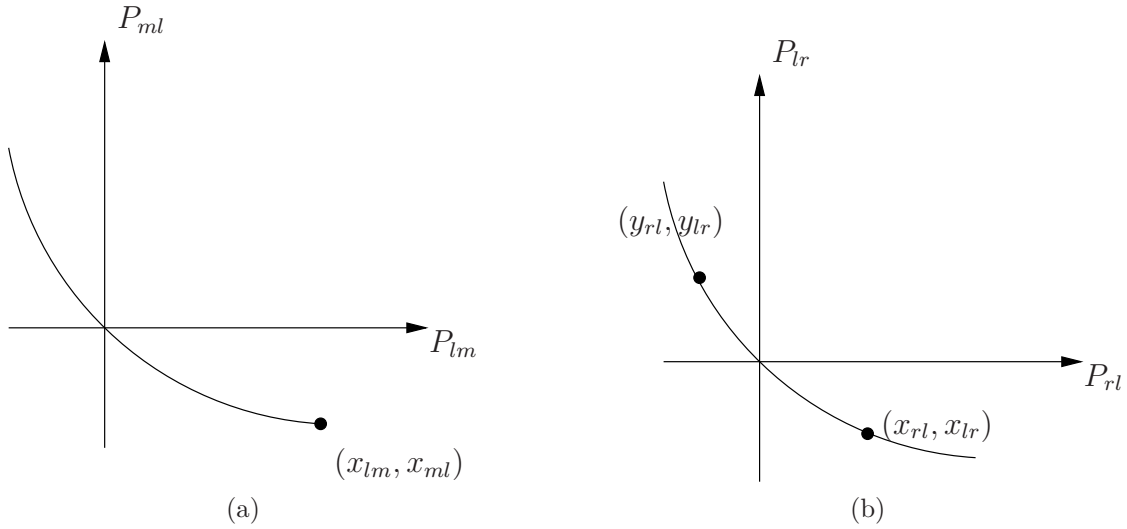


Figure A.2: Figure (a) shows the flow region for the line (l, m), where (x_{lm}, x_{ml}) lies at its lower right corner due to $c_l \leq 0$ and $c_m > 0$. Figure (b) shows the flow region for the line (r, l) to illustrate that $x_{rl} \geq y_{rl}$ (due to $\mathcal{F}_{\theta_{rl}} = \mathcal{O}(\text{conv}(\mathcal{F}_{\theta_{rl}}))$ and $y_{lr} \geq x_{lr}$).

corner of $\mathcal{F}_{\theta_{lm}}$. Thus, $x_{lm} \geq y_{ml}$ because of $(y_{lm}, y_{ml}) \in \mathcal{F}_{\theta_{lm}}$. Let $\mathcal{N}_l$ denote the set of all nodes connected to l that are neighbors of $\mathcal{G}$. One can derive the inequality $x_{lm} \geq y_{lm}$ for every $m \in \mathcal{N}_l$. Combining this set of inequalities with $x_l = \underline{P}_l \leq y_l$ or equivalently

$$x_l = \sum_{m \in \mathcal{N}_l} x_{lm} + x_{lr} \leq \sum_{m \in \mathcal{N}_l} y_{lm} + y_{lr} = y_l,$$

yields that $y_{lr} \geq x_{lr}$. As illustrated in Figure A.2b, this implies that $x_{rl} \geq y_{rl}$. This line of argument can be pursued until node 1 of the tree is reached. In particular, since node 1 is assumed to be a boundary node of $\mathcal{G}$, it can be shown by induction that $x_{1l} \geq y_{1l}$ for every node l such that $(1, l) \in \mathcal{E}$. On the other hand,

$$\sum_{l \in \mathcal{V}:\ l \sim 1} x_{1l} = x_1 = \underline{P}_1 \leq y_1 = \sum_{l \in \mathcal{V}:\ l \sim 1} y_{1l}$$

Therefore, the equality $x_{1l} = y_{1l}$ must hold for every $l \sim 1$. By propagating this equality down the subtree $\mathcal{G}$, we obtain that $x_{ik} = y_{ik}$ and $x_{ki} = y_{ki}$. This completes the proof of the relation $\mathcal{O}(\mathcal{S}) \subseteq \mathcal{O}(\mathcal{P})$.

In order to complete the proof of the lemma, it remains to show that $\mathcal{O}(\mathcal{P}) \subseteq \mathcal{O}(\mathcal{S})$. To this end, assume by contradiction that there is a point $\mathbf{p} \in \mathcal{O}(\mathcal{P})$ such that $\mathbf{p} \notin \mathcal{O}(\mathcal{S})$. In light of $\mathcal{P} \subseteq \mathcal{S}$, there exists $\mathbf{p}' \in \mathcal{O}(\mathcal{S})$ such that $\mathbf{p}' \leq \mathbf{p}$ with strict inequality in at least one coordinate. However, since $\mathcal{O}(\mathcal{S}) \subseteq \mathcal{O}(\mathcal{P})$, $\mathbf{p}'$ belongs to $\mathcal{P}$. This contradicts the assumption $\mathbf{p} \in \mathcal{O}(\mathcal{P})$.

A.3 Proof of Theorem 6

The proof proceeds along the same lines of Theorem 4. The requirement that $\underline{Q}_i < \beta_i$ for $i = 2, \ldots, n$ is to ensure that the reactive lower bound is in fact never tight for all the nodes in the network. Let h be parent of i and k be a child of i. Since we assume that power always flow from parent to child in the network, and from the angle constraints, $0 \leq \theta_{hi} \leq \tan^{-1}(\frac{b_{ik}}{g_{ik}})$. Over this range, $Q_{ih} \geq 0$. This corresponds to the intuition that reactive power should flow up the tree to support the voltage. Note the inductive line is very lossy in terms of reactive powers, therefore i might receive or supply reactive power to k. The Q_{ik} is monotonic in θ_{ik} starting at $\theta_{ik} = 0$ until it reaches is minimum at an angle of $\tan^{-1}(\frac{g_{ik}}{b_{ik}})$. Let $\tilde{\theta}_{ik} = \min(\tan^{-1}(\frac{g_{ik}}{b_{ik}}), \overline{\theta}_{ik})$, then $Q_i = Q_{ih} + \sum_{k:k\in\mathcal{C}(i)} Q_{ik} \geq \sum_{k:k\in\mathcal{C}(i)} Q_{ik} \geq \sum_{k:k\in\mathcal{C}(i)} b_{ik} - g_{ik}\sin(\tilde{\theta}_{ik}) - b_{ik}\cos(\theta(\theta)_{ik}) = \beta_i$. Therefore if $\underline{Q}_i < \beta_i$, the lower bounds on the reactive power injections are never tight.

Since by construction the lower bonds on the reactive power injection are never tight, we can ignore them from now on. Let $\mathcal{P}$ be the feasible region, that is, $\mathcal{P} = \{\mathbf{p} : \exists \mathbf{v} \in \mathbb{C}^n, P_i = \mathrm{Tr}(\mathbf{A}_i\mathbf{v}\mathbf{v}^H), |V_i| = 1, \underline{P}_i \leq P_i \leq \overline{P}_i, \mathrm{Tr}(\mathbf{B}_i\mathbf{v}\mathbf{v}^H) \leq \overline{Q}_i, |\theta_{ik}| < \overline{\theta}_{ik}, \forall i \sim k\}$. We can equivalently write $\mathcal{P}$ as $\mathcal{P} = \mathbf{A}(\mathcal{F}_\theta \cap \mathcal{F}_P \cap \mathcal{F}_Q)$, Since $\mathcal{F}_P$ and $\mathcal{F}_Q$ are defined by linear inequalities, they are convex. However, $\mathcal{F}_\theta$ is not.

Define $\mathcal{S} = \mathbf{A}(\mathrm{convhull}(\mathcal{F}_\theta) \cap \mathcal{F}_P \cap \mathcal{F}_Q)$. It is convex, and contains $\mathcal{P}$. To establish the theorem, it suffices to show the following lemma.

Lemma 17. *Suppose $\mathcal{P}$ is not empty, then $\mathcal{P} = \mathcal{O}(\mathcal{S})$.*

To prove the Lemma, let $\mathbf{p}^* \in \mathcal{S}$ be the optimal solution of the relaxed problem, $\mathbf{f}^* \in \mathrm{convhull}(\mathcal{F}_\theta) \cap \mathcal{F}_P \cap \mathcal{F}_Q$ its corresponding active flow vector and $\mathbf{r}^* = \mathbf{H}\mathbf{f}^*$ be the corresponding reactive power flow vector. It suffices to show that if $P_i^* > \underline{P}_i$, then $(f_{ik}^*, f_{ki}^*) \in \mathcal{F}_{\theta,ik}$ for every $k \sim i$. Once this fact is established, the rest of the proof is the same as the proof of Lemma 5. Suppose that $P_i^* > \underline{P}_i$, but $(f_{ik}^*, f_{ki}^*) \notin \mathcal{F}_{\theta,ik}$ for some k. Then there exists $\epsilon > 0$ such that $(f_{ik}^* - \epsilon, f_{ki}^*) \in \mathrm{convhull}(\mathcal{F}_{\theta,ik})$. Let $(\tilde{f}_{ik}, \tilde{f}_{ki}) = (f_{ik}^* - \epsilon, f_{ki}^*)$. Since

$$\mathbf{H}_{ik}\begin{bmatrix}-\epsilon\\0\end{bmatrix} = -\frac{\epsilon}{2b_{ik}g_{ik}}\begin{bmatrix}b_{ik}^2 - g_{ik}^2\\ b_{ik}^2 + g_{ik}^2\end{bmatrix} < \mathbf{r}^*,$$

Therefore $(\tilde{f}_{ik}, \tilde{f}_{ki})$ is a better feasible flow on the line (i,k), which contradicts the optimality of $\mathbf{f}^*$.

A.4 Proof of Lemma 7

The proofs provided for the case with fixed voltage magnitudes can be easily adapted to prove (2.33a) and (2.33b). Therefore, we only prove (2.33c) here. It follows from (2.32) that there exists a linear transformation from $\mathcal{H}_{ik}(\tilde{\mathbf{v}})$ to $\overline{\text{conv}}(\mathcal{F}_{\theta_{ik}}(\tilde{\mathbf{v}}))$, which is independent of $\tilde{\mathbf{v}}$. We denote this transformation as $\overline{\text{conv}}(\mathcal{F}_{\theta_{ik}}(\tilde{\mathbf{v}})) = l_{ik}(\mathcal{H}_{ik}(\tilde{\mathbf{v}}))$. Define $\mathcal{H}(\tilde{\mathbf{v}})$ as $\bigcap_{(i,k)\in\mathcal{E}} \mathcal{H}_{ik}(\tilde{\mathbf{v}})$ and the function $l(\cdot)$ as the natural extension of l_{ik}. Hence, $\overline{\text{conv}}(\mathcal{F}_\theta(\tilde{\mathbf{v}})) = l(\mathcal{H}(\tilde{\mathbf{v}}))$. One can write:

$$\begin{aligned}\bigcup_{\underline{\mathbf{v}}\le\tilde{\mathbf{v}}\le\overline{\mathbf{v}}} \overline{\text{conv}}(\mathcal{P}_\theta(\tilde{\mathbf{v}})) &= \bigcup_{\underline{\mathbf{v}}\le\tilde{\mathbf{v}}\le\overline{\mathbf{v}}} \mathbf{A}\overline{\text{conv}}(\mathcal{F}_\theta(\tilde{\mathbf{v}})) \\ &= \mathbf{A} \bigcup_{\underline{\mathbf{v}}\le\tilde{\mathbf{v}}\le\overline{\mathbf{v}}} l(\mathcal{H}(\tilde{\mathbf{v}})) \\ &= \mathbf{A}l\Big(\bigcup_{\underline{\mathbf{v}}\le\tilde{\mathbf{v}}\le\overline{\mathbf{v}}} \mathcal{H}(\tilde{\mathbf{v}}) \Big)\end{aligned} \tag{A.12}$$

On the other hand, $\bigcup_{\underline{\mathbf{v}}\le\tilde{\mathbf{v}}\le\overline{\mathbf{v}}} \mathcal{H}(\tilde{\mathbf{v}})$ is a convex set because it consists of all Hermitian matrices $\mathbf{W}$ whose entries satisfy certain linear and convex constraints. Due to the convexity of this set as well as the linearity of $\mathbf{A}$ and l, it can be concluded from (A.12) that $\bigcup_{\underline{\mathbf{v}}\le\tilde{\mathbf{v}}\le\overline{\mathbf{v}}} \overline{\text{conv}}(\mathcal{P}_\theta(\tilde{\mathbf{v}}))$ is convex. ■

Appendix B

Proofs of Results in Chapter 4

B.1 Proof of Theorem 14

The proof of Theorem 14 mainly follows duality of the DC-OPF problem. The statement that the two step strategy is optimal is not surprising, since the small-σ assumption guarantees that the generation pattern (the sign of g_i at bus i) and the congestion pattern (which lines are congested) are the same at day-ahead and real-time. The interesting part is to show that the size of the perturbed problem can be reduced to a problem of size $K+1$. The rest of the proof shows that each step of Algorithm 1 is valid.

Let $\overline{\mathbf{g}}$ be the solution of $J(\boldsymbol{\alpha}, \hat{\mathbf{d}})$. Let the lines be numbered as in step 2 of Algorithm 1. Lemma 18 states that the total number of buses with + label can be no greater than $K+1$.

Lemma 18. *Let N_+ be the number of buses labelled with +. Then $N_+ \leq K+1$. Furthermore, solution of $\mathbf{A}_2\boldsymbol{\lambda} = \mathbf{0}$ can be expressed as a linear combination of λ_i where i is labelled with +, together $K+1-|N_+|$ other λ's.*

Suppose Lemma 18 is true and renumber the buses as in step 5 of Algorithm 1. For a generic set of capacities, all congested flows are fundamental. Since only the first K lines are congested, $J(\boldsymbol{\alpha}, \hat{\mathbf{d}})$ can be equivalently written as

$$J(\boldsymbol{\alpha}, \hat{\mathbf{d}}) = \min_{\overline{\mathbf{g}}} \ \alpha^T (\overline{g})^+ \tag{B.1a}$$

$$\text{s.t. } \overline{g} - \mathbf{A}\overline{\mathbf{f}} - \hat{\mathbf{d}} = \mathbf{0} \tag{B.1b}$$

$$\overline{f}_k = c_k, \ k = 1, \ldots, K. \tag{B.1c}$$

The dual is

$$\max \overline{\boldsymbol{\lambda}}^T \hat{\mathbf{d}} - \sum_{k=1}^{K} \overline{\mu}_k c_k \tag{B.2a}$$

$$\text{s.t. } \mathbf{0} \leq \overline{\boldsymbol{\lambda}} \leq \boldsymbol{\alpha} \tag{B.2b}$$

$$\mathbf{A}^T \overline{\boldsymbol{\lambda}} + \begin{bmatrix} \overline{\mu}_1 & \ldots & \overline{\mu}_K & 0 & \ldots & 0 \end{bmatrix}^T = \mathbf{0} \tag{B.2c}$$

where $\overline{\boldsymbol{\lambda}}$ are the Lagrange multipliers to the energy balance constraint (B.1b) and $\overline{\boldsymbol{\mu}}$ are the Lagrange multipliers to the flow constraints (B.1c). Part of the vector constraint in (B.2c) can be written as $\mathbf{A}_2\overline{\boldsymbol{\lambda}} = \mathbf{0}$ where $\mathbf{A}_2$ is defined in Algorithm 1. By the choice of $\lambda_1, \ldots, \lambda_{K+1}$, $\overline{\lambda}_{K+1}, \ldots, \overline{\lambda}_n$ can be written as a linear combination of $\overline{\lambda}_1$ to $\overline{\lambda}_{K+1}$. This fact is used later for the dual of the second stage problem as well.

Now write the solution of the original problem as $\overline{g} + \boldsymbol{\Delta}$. The objective function becomes

$$\alpha^T(\overline{g} + \boldsymbol{\Delta})^+.$$

And denote the flows by $\overline{\mathbf{f}} + \mathbf{f}$. Using the small-$\sigma$ assumption, if $\overline{g}_i < 0$, then Δ_i is small enough such that $\overline{g}_i + \Delta_i < 0$ always and the price of Δ_i can be thought as 0; if $\overline{g}_i > 0$, then $\overline{g}_i + \Delta_i$ always and there are no constraint on Δ_i. The power balance constraints in the second stage becomes

$$\mathbf{y} + \overline{g} + \boldsymbol{\Delta} - \mathbf{A}(\overline{\mathbf{f}} + \mathbf{f}) - (\hat{\mathbf{d}} + \mathbf{e}) = \mathbf{0},$$

where $\mathbf{y}$ is the energy purchased at the second stage. By definition of $\overline{g}$ and $\overline{\mathbf{f}}$, this constraint becomes

$$\mathbf{y} + \boldsymbol{\Delta} - \mathbf{A}\mathbf{f} - \mathbf{e} = \mathbf{0}.$$

The flow limits are

$$\overline{f}_k + f_k \leq c_k,\ k = 1, \ldots, K$$

and by the small-σ assumption and $\overline{f}_k = c_k$, they become

$$f_k \leq 0,\ k = 1, \ldots, K.$$

Therefore under the small-σ assumption, the two stage optimization problem becomes

$$\begin{aligned} \min_{\boldsymbol{\Delta}} \boldsymbol{\alpha'}^T\Delta + \mathbb{E}[\min_{\mathbf{y}} \boldsymbol{\beta'}(\mathbf{y})^+\}] \\ \text{s.t. } \mathbf{y} + \boldsymbol{\Delta} - \mathbf{A}\mathbf{f} - \mathbf{e} = \mathbf{0} \\ f_k \leq 0\ k = 1, \ldots, K, \end{aligned} \tag{B.3}$$

where the constraint $\Delta_i = 0$ if $\overline{g}_i = 0$. Recall the $\alpha'_i = 0$ ($\beta'_i = 0$) if bus i is labelled $-$ and α_i (β) otherwise. This means that under the small-σ assumption, we can assume that a bus dumping energy in the nominal problem as a source of energy with cost 0.

The dual of the second stage problem in (B.3) is

$$\max \boldsymbol{\lambda}^T(\mathbf{e} - \boldsymbol{\Delta}) \tag{B.4a}$$

$$\text{s.t. } \mathbf{0} \leq \boldsymbol{\lambda} \leq \boldsymbol{\beta'} \tag{B.4b}$$

$$\mathbf{A}^T\boldsymbol{\lambda} + \begin{bmatrix}\mu_1 & \ldots & \mu_K & 0 & \ldots & 0\end{bmatrix}^T = \mathbf{0} \tag{B.4c}$$

$$\mu_k \geq 0,\ k = 1, \ldots, K. \tag{B.4d}$$

This is exactly the form of (B.2). Therefore we can express λ_{K+2}, λ_n as a function of $\lambda_1, \ldots, \lambda_{K+1}$ through the $\mathbf{B}$ matrix defined in step 4 of Algorithm 1. Let $\tilde{\boldsymbol{\lambda}} = [\lambda_1 \ \ldots \ \lambda_{K+1}]^T$ and write (B.4) in terms of $\tilde{\boldsymbol{\lambda}}$ gives

$$\max \ \tilde{\boldsymbol{\lambda}}^T(\tilde{\mathbf{e}} - \tilde{\boldsymbol{\Delta}}) \tag{B.5a}$$

$$\text{s.t. } \mathbf{0} \leq \tilde{\boldsymbol{\lambda}} \leq \tilde{\boldsymbol{\beta}} \tag{B.5b}$$

$$\tilde{\mathbf{A}}\tilde{\boldsymbol{\lambda}} \leq \mathbf{0}, \tag{B.5c}$$

where the definitions for $\tilde{\mathbf{A}}$, $\tilde{\mathbf{e}}$, $\tilde{\boldsymbol{\beta}}$ and $\tilde{\boldsymbol{\Delta}}$ are given in Algorithm 1. Since each column of B sum up to 1, and β_i' is either 0 or β, it suffices to consider the first $K+1$ as in (B.5b). Taking the dual of (B.5), we get back the primal problem

$$\tilde{J}(\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\Delta}}) = \min_{\mathbf{y}} \ \tilde{\boldsymbol{\alpha}}^T(\mathbf{y})^+ \tag{B.6a}$$

$$\text{s.t. } \mathbf{y} - \tilde{A}\tilde{\mathbf{f}} - (\tilde{\mathbf{e}} - \tilde{\boldsymbol{\Delta}}) = \mathbf{0} \tag{B.6b}$$

$$\tilde{\mathbf{f}} \leq 0, \tag{B.6c}$$

but this is exactly the primal of a second stage problem with a network of size $K+1$, with flows governed by the matrix $\tilde{\mathbf{A}}$. Therefore the two stage problem is

$$\min_{\tilde{\boldsymbol{\Delta}}} \tilde{\boldsymbol{\alpha}}^T \tilde{\boldsymbol{\Delta}} + \mathbb{E}[\tilde{J}(\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\Delta}}], \tag{B.7}$$

and this is precisely the risk limiting dispatch problem for a network of size $K+1$.

B.1.1 Proof of Lemma 18

Since $\mathbf{A}_2$ has full rank, by the rank-nullity theorem, the null space of $\mathbf{A}_2$ has rank $K+1$. Therefore all of λ_i can be expressed as linear combination of some $K+1$ λ's. By complementary slackness [25], the buses with label + ($\overline{g}_i > 0$) has $\lambda_i = \alpha_i$. Intuitively, this means buses that are generating positive power are generating at their marginal costs. Since these are the marginal generators, they cannot dependent on other λ_i's [73]. Therefore any set of independent λ_i's must include the buses labelled +.

Appendix C

Optimal Power Flow for Lossless Networks

The results in this section is different from Chapter 2 in three aspects

- We focus on lossless networks.
- Only voltage constraints are considered.
- We look at the convex hull instead of the Pareto-front.

Therefore the results in this section are of a weaker flavor since we need to assume that the networks are lossless and we only consider voltage constraints. They are useful since in practice some distribution networks consists of a ring feeder and trees hanging of the feeder nodes as in Figure C.1. In this case, the objective functions are often to minimize the loss

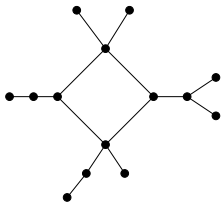


Figure C.1: A distribution network with a ring feeder.

at the feeders. Also, the feeder nodes are generally considered as slack buses, so they only have a voltage constraint. Since minimizing a linear function over $\mathcal{A}$ and convhull($\mathcal{A}$) has the same objective values, characterizing the convex hull of the injection region is useful.

The voltage constraint injection region is defined as

$$\mathcal{P} = \{\mathbf{p} : \mathbf{p} = \mathrm{Re}(\mathrm{diag}(\mathbf{v}\mathbf{v}^H\mathbf{Y}^H)), \underline{V}_i \leq |V|_i \leq \overline{V}_i\}. \tag{C.1}$$

We can again define a enlarged convex region $\tilde{\mathcal{P}}$ as

$$\tilde{\mathcal{P}} = \{\mathbf{p} : \mathbf{p} = \mathrm{Re}(\mathrm{diag}(\mathbf{W}\mathbf{Y}^H)), \underline{V}_i^2 \leq W_{ii} \leq \overline{V}_i^2, \mathbf{W} \succcurlyeq 0\}. \tag{C.2}$$

We have the following theorem

Theorem 19. *Given a network with n buses represented by its bus admittance matrix* $\mathbf{Y}$*. Let* $\mathcal{P}$ *and* $\tilde{\mathcal{P}}$ *be defined as in* (C.1) *and* (C.2) *respectively. Then if the network is a* lossless cycle *or a* lossless cycle with one chord*, then* convhull$(\mathcal{P}) = \tilde{\mathcal{P}}$*.*

The next theorem states that joining the basic types of networks in a certain way preserves the characterization result. Given two networks G and H, the network K is said to be a 1-connection of G and H if it is possible to decompose K into two components K_1 and K_2 such that they have only one node in common and no edges between them, where K_1 is equal to G and K_2 is equal to H. Note by equal we mean that the admittance matrices are identical. In particular, if a line in G or H is lossless then its corresponding line in K is also lossless. We say K is obtained by 1-connecting G and H. Figure C.1 gives an example of a network obtained by 1-connecting a cycle and a number of trees.

Theorem 20. *Given a network on n nodes with voltage constraints. Then* convhull $\mathcal{P} = \tilde{\mathcal{P}}$ *if the network is a result of repeatedly 1-connecting a lossless cycle and a tree.*

It is simple to check if a network has the topology that satisfies the conditions in Theorem 20. Given a network, first decompose it into its one connected parts which can be done in linear time. Then one simply check each of the parts to see if they are a tree or a lossless cycle.

First we prove Theorem 19. This requires that we prove an analogous result about trees first. Consider the following lemma

Lemma 21. *Given a* tree *network with n buses. Let* $\mathcal{P}$ *and* $\mathcal{P}_{\mathbf{W}}$ *be defined as in* (C.1) *and* (C.2) *respectively. Then* convhull$(\mathcal{P}) = \mathcal{P}_W$*.*

Proof. To prove this theorem, it suffices to prove that minimizing linear functions over $\mathcal{P}$ and $\mathcal{P}_{\mathbf{W}}$ has the same optimal objective value for all coefficients [25]. So consider the optimization problem

$$\begin{aligned} J = \text{minimize } & \sum_{i=1}^{n} c_i P_i \\ \text{subject to } & \underline{V}_i \leq |V_i| \leq \overline{V}_i, \ \forall i \\ & \mathbf{p} = \mathrm{Re}(\mathrm{diag}(\mathbf{v}\mathbf{v}^H\mathbf{Y}^H)), \end{aligned} \tag{C.3}$$

and its relaxation

$$\begin{aligned} J_1 = \text{minimize } & \sum_{i=1}^{n} c_i P_i \\ \text{subject to } & \underline{V}_i^2 \le W_{ii} \le \overline{V}_i^2, \ \forall i \\ & \mathbf{W} \succcurlyeq 0 \\ & \mathbf{p} = \text{Re}(\text{diag}(\mathbf{W}\mathbf{Y}^H)), \end{aligned} \tag{C.4}$$

where the costs can be general (no longer constraint to be positive). We show $J = J_1$ for all c_i's.

The dual of (C.4) is

$$\begin{aligned} J_1 = \text{maximize } & \sum_{i=1}^{n} (\underline{V}_i^2 \underline{\lambda}_i - \overline{V}_i^2 \overline{\lambda}_i) \\ \text{subject to } & \mathbf{\Lambda} + \mathbf{M} \succcurlyeq 0. \end{aligned} \tag{C.5}$$

From Theorem 13 if the costs are such that $\mathbf{M}$ is connected, then the optimal solution to (C.4) is rank 1 and clearly $J = J_1$. If $\mathbf{M}$ is disconnected, then $\mathbf{M}$ can be written as a block diagonal matrix. Suppose there are K connected components of $\mathbf{M}$, then $\mathbf{M} = \text{diag}(\mathbf{M}_1, \ldots, \mathbf{M}_K)$. Since the network is a tree, $\mathbf{M}_i$ fits the topology of a tree for each i. Then (C.4) and (C.5) decomposes into K independent primal-dual subproblems, and we may apply Theorem 13 to each of them. Let $W_1^*, \ldots, W_K^*$ denote the optimal solutions to each of the subproblems. By Theorem 13, they are all rank 1 so we can write $W_i^* = \mathbf{v}_i^* (\mathbf{v}_i^*)^H$ for each i. An optimal solution $\mathbf{W}^*$ to the original problem is given by $\mathbf{W}^* = \mathbf{v}^* (\mathbf{v}^*)^H$ where $\mathbf{v}^* = \begin{bmatrix} \mathbf{v}_1^* \\ \vdots \\ \mathbf{v}_K^* \end{bmatrix}$. □

Now we prove Theorem 19. The approach is the same as in the proof of Lemma 21. That is, we look at (C.3), (C.4) and their dual (C.5). We say a matrix $\mathbf{A}$ is lossless if all the off diagonal terms of A are purely imaginary or 0. We prove the following lemma

Lemma 22. *Given a graph on n nodes that is either an odd cycle or a cycle with one chord, if $\mathbf{A}$ is lossless, positive semidefinite and fits G, then* $\text{rank}(\mathbf{A}) \ge n - 1$.

Theorem 19 can be proved from Lemma 22. Suppose the electrical network is lossless and has the topology of an odd cycle or a cycle with one chord. The network being lossless means $\mathbf{Y}$ is purely imaginary, and $\mathbf{M} = \frac{1}{2}(\mathbf{C}\mathbf{Y} + \mathbf{Y}^H\mathbf{C})$ is also purely imaginary since $\mathbf{C}$ is real. Suppose that the costs are such that $M_{ik} \neq 0$ if (i, k) is connected by a line in the network. Since $\mathbf{\Lambda}^*$ is diagonal, the dual matrix $\mathbf{\Lambda}^* + \mathbf{M}$ is positive semidefinite, lossless and fits the network topology. Apply Lemma 22 shows $\mathbf{W}^*$ is rank 1. If the cycle is even, we add a chord between two buses, and let the admittance of that chord go to 0. Since

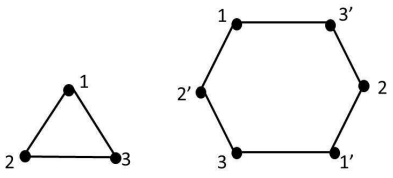
all the functions in (C.4) are continuous, the optimal solution of the network with a chord approaches the network without the chord as the admittance goes to 0.

If the costs are such that $M_{ik} = 0$ even if (i, k) is connected in the network, then $\mathbf{M}$ either fits a tree or becomes disconnected. If $\mathbf{M}$ fits a tree, then apply Theorem 13. If $\mathbf{M}$ becomes disconnected, then $\mathbf{M}$ can be written as a block diagonal matrix. If there are K connected components of $\mathbf{M}$, then $\mathbf{M} = \text{diag}(\mathbf{M}_1, \ldots, \mathbf{M}_K)$. Since the network is a cycle (with a chord), then $\mathbf{M}_i$ is either a tree or a cycle for each i. We can apply Theorem 13 or Lemma 22 to each component and obtain an optimal solution $\mathbf{W}^*$ in the same way as in the tree network case. To finish Theorem 19, it remains to proof Lemma 22.

Proof. Given a graph G, the tree-width of G is a number that intuitively captures how close G is to a tree. For example, the tree-width of a tree is 1, and the tree-width of a cycle is 2. The rigorous definition and some methods of computing the tree-width the reader may consult [77]. A graph of tree-width 2 is also called serial-parallel graph or a partial-2-tree. The following lemma collects the known results that we need.

Lemma 23. *If G is a cycle of length n, then the minimum rank of real positive semidefinite matrices fitting G is $n - 2$ [78]. More generally, if the graph has tree-width 2, the minimum rank is $n - 2$ [79, 80].*

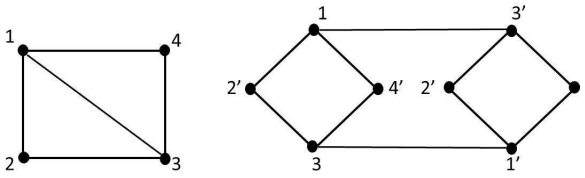
Given a graph G with n nodes and m edges. We construct a bipartite graph derived from G that we call the bipartite expansion of G and denote by $B(G)$. $B(G)$ is a bipartite graph with $2n$ nodes and $2m$ edges. Label the nodes $1, 2, \ldots, n, 1', 2', \ldots, n'$ with the bipartition being $\{1, \ldots, n\}$ and $\{1', \ldots, n'\}$. There is an edge between i and k' if and only if $i \neq k$ and (i, k) is an edge in G. If G is an odd cycle then $B(G)$ is also a cycle and if G is a cycle with a chord then $B(G)$ has tree-width 2 (a subclass of linear-2-trees in the language of [80]). Two examples are given in Figures C.2 and C.3. If G is an even cycle then $B(G)$ is two disconnected cycles, therefore the assumption of odd cycle is needed in the Lemma.



(a) G (b) $B(G)$

Figure C.2: (a) shows a 3-cycle and (b) shows its bipartite expansion.

Given a graph G, suppose $\mathbf{A}$ is lossless, positive semidefinite and $\mathbf{A}$ fits G. We show that the rank of $\mathbf{A}$ cannot be lower than $n - 1$. Suppose $\mathbf{A}$ has rank r. Then $\mathbf{A}$ can be factored as $\mathbf{A} = \mathbf{Z}^H\mathbf{Z}$ for some complex matrix $r \times n$ matrix $\mathbf{Z}$. Let $\mathbf{z}_1, \ldots, \mathbf{z}_n \in \mathbb{C}^r$ be the columns



(a) G (b) $B(G)$

Figure C.3: (a) shows a 4-cycle with a chord and (b) shows its bipartite expansion.

of $\mathbf{Z}$. They satisfy the graph topology condition

$$\mathbf{z}_i^H \mathbf{z}_k = \begin{cases} 0 & \text{if } i \nsim k \\ \neq 0 & \text{if } i \sim k \end{cases} \tag{C.6}$$

and the lossless line condition

$$\mathrm{Re}(\mathbf{z}_i^H \mathbf{z}_k) = 0 \text{ if } i \neq k. \tag{C.7}$$

From each complex vector we define two real vectors as

$$\mathbf{x}_i = \begin{bmatrix} \mathrm{Re}(\mathbf{z}_i) \\ \mathrm{Im}(\mathbf{z}_i) \end{bmatrix} \qquad \mathbf{y}_i = \begin{bmatrix} \mathrm{Im}(\mathbf{z}_i) \\ -\mathrm{Re}(\mathbf{z}_i) \end{bmatrix}$$

Since $\mathbf{z}_i \in \mathbb{C}^r$, then $\mathbf{x}_i, \mathbf{y}_i \in \mathbb{R}^{2r}$. By algebra, $\mathrm{Re}(\mathbf{z}_i^H \mathbf{z}_k) = \mathbf{x}_i^T \mathbf{x}_k = \mathbf{y}_i^t \mathbf{y}_k$ and $\mathrm{Im}(\mathbf{z}_i^H \mathbf{z}_k) = \mathbf{x}_i^T \mathbf{y}_k$. In terms of $\mathbf{x}$'s and $\mathbf{y}$'s, (C.6) becomes

$$\mathbf{x}_i^T \mathbf{y}_k = \begin{cases} 0 & \text{if } i \nsim k \\ \neq 0 & \text{if } i \sim k \end{cases} \tag{C.8}$$

and (C.7) becomes

$$\mathbf{x}_i^T \mathbf{x}_k = \mathbf{y}_i^T \mathbf{y}_k = 0 \text{ if } i \neq k. \tag{C.9}$$

Define the matrix B to be the $2r \times 2n$ matrix with columns $\mathbf{x}_1, \ldots, \mathbf{x}_n, \mathbf{y}_1, \ldots, \mathbf{y}_n$. By (C.8) and (C.9) B fits $B(G)$. But if G is an odd cycle or a cycle with one chord, applying Lemma 23 to $B(G)$ gives $\mathrm{rank}(B) \geq 2n - 2$. Thus $2r \geq 2n - 2$ or $r \geq n - 1$. □

Now we proceed to the proof Theorem 20. Given a network G, we say the matrix $\mathbf{A}$ satisfies G if $\mathbf{A}$ fits the topology of G and A_{ik} is purely imaginary if the line from bus i to bus k is lossless. We have the following lemma.

Lemma 24. *Given two networks G and H with n and m buses respectively, let K be a network obtained by 1-connecting G and H, so K has $n + m - 1$ buses. If $\mathbf{A}$ is a positive semidefinite matrix that satisfies K, then* $\mathrm{rank}(\mathbf{A}) \geq n + m - 2$.

From the basic topologies in Theorem 19, we can apply the Lemma 24 repeatedly to get Theorem 20. A version of Lemma 24 just about graphs (without considering lossless lines and such) is known in the graph theory community [78, 81]. We give a proof here to show the additional condition of lossless lines does not change the result.

Proof. Let G, H and K be networks given in the statement of the Lemma. Label the buses in K to be $1, 2, \ldots, n-1, n, n+1, n+2, \ldots, n+m-1$ where the subnetwork induced by $1, \ldots, n-1, n$ corresponds to G and the subnetwork induced by $n, n+1, n+m-1$ corresponds to H. So bus n is the common bus in the 1-connection. Suppose $\mathbf{A}$ is a $(n+m-1) \times (n+m-1)$ positive semidefinite matrix that satisfies K and has rank r. Then it is possible to factor $\mathbf{A}$ as $\mathbf{A} = \mathbf{Z}^H\mathbf{Z}$ for some $r \times (n+m-1)$ matrix $\mathbf{Z}$. Let $\mathbf{z}_1, \ldots, \mathbf{z}_{n+m-1}$ be the columns of $\mathbf{Z}$. Let $\mathcal{U}$ be the subspace spanned by $\mathbf{z}_1, \ldots, \mathbf{z}_{n-1}$ and $\mathcal{V}$ be the subspace spanned by $\mathbf{z}_{n+1}, \ldots, \mathbf{z}_{n+m-1}$. By construction of K, there are no lines between the set of buses $\{1, \ldots, n-1\}$ and $\{n+1, \ldots, n+m-1\}$. Therefore $\mathcal{V}$ is orthogonal to $\mathcal{U}$. We may write vector $\mathbf{z}_n$ as $\mathbf{z}_n = \mathbf{u} + \mathbf{v} + \mathbf{w}$ where $\mathbf{u} \in \mathcal{U}$, $\mathbf{v} \in \mathcal{V}$ and $\mathbf{w}$ is orthogonal to $\mathcal{U}$ and $\mathcal{V}$. Let $\mathbf{Z}_G$ be the matrix with columns $\mathbf{z}_1, \ldots, \mathbf{z}_{n-1}, \mathbf{u}$ and $\mathbf{Z}_H$ be the matrix with columns $\mathbf{v}, \mathbf{z}_{n+1}, \ldots, \mathbf{z}_{n+m-1}$. Let $\mathbf{A}_G = \mathbf{Z}_G^H\mathbf{Z}_G$. Since $\mathbf{z}_i^H\mathbf{u} = \mathbf{z}_i^H\mathbf{z}_n$ for $i = 1, \ldots, n-1$, $\mathbf{A}_G$ equals the matrix formed by the first n rows and n columns of $\mathbf{A}$. By the assumption $\mathbf{A}$ satisfies K, so $\mathbf{A}_G$ satisfies G. Similarly $\mathbf{Z}_H^H\mathbf{Z}_H$ satisfies H. By the assumption in the Lemma, we have $\operatorname{rank}(\mathbf{Z}_G) \geq n-1$ and $\operatorname{rank}(\mathbf{Z}_H) \geq m-1$, so equivalently $\dim\mathcal{U} \geq n-1$ and $\dim\mathcal{V} \geq m-1$. Since $\mathcal{U}$ is orthogonal to $\mathcal{V}$ and $\mathbf{z}_1, \ldots, \mathbf{z}_{n+m-1}$ spans $\mathcal{U} + \mathcal{V}$, $\operatorname{rank}(\mathbf{A}) = \dim\mathcal{U} + \dim\mathcal{V} \geq (n-1) + (m-1) = n+m-2$. □