

# **Optimization Methods for Power Grid Reliability**

**Sean Harnett**

Submitted in partial fulfillment of the  
requirements for the degree  
of Doctor of Philosophy  
in the Graduate School of Arts and Sciences

**COLUMBIA UNIVERSITY**

2016

©2016

Sean Harnett

All Rights Reserved

# **ABSTRACT**

## **Optimization Methods for Power Grid Reliability**

**Sean Harnett**

This dissertation focuses on two specific problems related to the reliability of the modern power grid. The first part investigates the economic dispatch problem with uncertain power sources. The classic economic dispatch problem seeks generator power output levels that meet demand most efficiently; we add risk-awareness to this by explicitly modeling the uncertainty of intermittent power sources using chance-constrained optimization and incorporating the chance constraints into the standard optimal power flow framework. The result is a dispatch of power which is substantially more robust to random fluctuations with only a small increase in economic cost. Furthermore, it uses an algorithm which is only moderately slower than the conventional practice.

The second part investigates “the power grid attack problem”: aiming to maximize disruption to the grid, how should an attacker distribute a budget of “damage” across the power lines? We formulate it as a continuous problem, which bypasses the combinatorial explosion of a discrete formulation and allows for interesting attacks containing lines that are only partially damaged rather than completely removed. The result of our solution to the attack problem can provide helpful information to grid planners seeking to improve the resilience of the power grid to outages and disturbances. Both parts of this dissertation include extensive experimental results on a number of cases, including many realistic large-scale instances.

# Contents

- List of Figures v
  
- List of Tables ix
  
- 1 Introduction 1**
  
- 2 Power grid preliminaries 3**
  - 2.1 Structure of power grids . . . . . 3
  - 2.2 Power grid stability requirements . . . . . 4
    - 2.2.1 Frequency stability . . . . . 4
    - 2.2.2 Voltage stability . . . . . 5
    - 2.2.3 Line stability . . . . . 5
  - 2.3 The power balance equations . . . . . 6
  - 2.4 The power flow problem . . . . . 9
    - 2.4.1 Existence and uniqueness of solutions . . . . . 10
    - 2.4.2 DC approximation . . . . . 10
  - 2.5 Power flow examples . . . . . 11
    - 2.5.1 Power circles . . . . . 11
    - 2.5.2 Nose curves . . . . . 15
    - 2.5.3 No solution . . . . . 16
    - 2.5.4 Multiple solutions . . . . . 16

2.6	Optimal power flow problem . . . . .	18
2.6.1	Difficulties solving the OPF . . . . .	19
2.6.2	Semidefinite relaxation of the OPF . . . . .	20
2.7	Difficult optimal power flow examples . . . . .	23
2.7.1	Two-bus OPF with local solutions . . . . .	24
2.7.2	Different solvers converge to different solutions . . . . .	24
2.7.3	Failure to converge after many iterations . . . . .	26
<b>I</b>	<b>The economic dispatch problem with uncertain power sources</b>	<b>27</b>
<b>3</b>	<b>Introduction</b>	<b>28</b>
3.1	Transmission system controls . . . . .	31
3.2	Standard generation dispatch . . . . .	32
3.3	Adjusting to power fluctuations . . . . .	34
3.4	Using chance constraints . . . . .	36
3.5	Uncertain power sources . . . . .	38
3.6	Affine Control . . . . .	39
3.7	Overview of the CC-OPF solution methodology . . . . .	40
<b>4</b>	<b>Solving the chance-constrained OPF problem</b>	<b>42</b>
4.1	Chance-constrained optimal power flow: formal expression . . . . .	42
4.1.1	Preliminaries . . . . .	43
4.1.2	Formal expression . . . . .	45
4.1.3	Objective . . . . .	46
4.1.4	Variance of the line flow $f_{km}$ . . . . .	47
4.2	Formulating the chance-constrained problem as a conic program . . . . .	48
4.3	Solving the conic program . . . . .	51
4.3.1	Example of typical performance of the cutting plane algorithm . . . . .	54

<b>5</b>	<b>Numerical experiments</b>	<b>56</b>
5.1	Failure of standard OPF	57
5.2	Cost of reliability under high wind penetration	58
5.3	Non-locality	60
5.4	Increasing penetration	61
5.5	Out-of-sample tests	61
5.6	Scalability	65
5.7	Changing locations for wind farms	66
5.8	Reversal of line flows	66
<b>II</b>	<b>The power grid attack problem</b>	<b>68</b>
<b>6</b>	<b>Introduction</b>	<b>69</b>
6.1	Motivation	69
6.1.1	Continuous formulation	70
6.1.2	Is this realistic?	71
6.2	Attacker budget and the upper-level problem	72
6.3	Optimal power flow lower-level problem	73
6.4	Related work	76
<b>7</b>	<b>Solving the attack problem</b>	<b>78</b>
7.1	Solving the OPF lower-level problem	79
7.2	Non-convexity and multiple local optima	79
7.3	The Frank-Wolfe algorithm	80
7.3.1	Example	83
7.4	Computing the gradient	83
7.5	Other approaches	84
7.5.1	Derivative-free optimization	84

7.5.2	Branching	85
<b>8</b>	<b>Numerical experiments</b>	<b>87</b>
8.1	9-bus example	87
8.1.1	Attack parameters	88
8.1.2	Results	88
8.1.3	Summary	89
8.2	57-bus example: branching	90
8.3	118-bus example: head and tail	94
8.4	Large-scale example on the Polish grid	95
	<b>Bibliography</b>	<b>98</b>

# List of Figures

- 2.1 Two-bus and three-bus grids for the examples in section 2.5. . . . . 12
- 2.2 Sending-end and receiving-end circles from [17]. Both circles have radius  $|B| = |y||V_1||V_2|$ . . . . . 14
- 2.3 Nose curves for two bus example from [29] with  $b = -10$ . Notice how as  $P_2$  increases from zero there are first two solutions, then one solution, and finally no solutions for each load factor. . . . . 16
- 2.4 Three bus example from [29] with  $|V_1| = 0$  and  $\theta_1 = 0$ . The black dots are the four possible solutions described in table 2.1. Notice in the right image the bands of red and green near the blue region which converge to solutions 2 and 3 rather than solution 1. . . . . 18
- 2.5 Two bus grid from [22]. Note that we use the impedance  $z$  rather than the admittance  $y$  in this example. . . . . 24
- 2.6 Two bus example from [22]. Note that the set of feasible points  $(V_1, V_2)$  is disconnected. Searching locally along the blue feasible region yields the non-optimal local solution at  $(.95, .976)$  marked with the dot. Searching in the red feasible region gives the global minimum marked by the star at  $(.952, 1.05)$ . . . 25

3.1	Bonneville Power Administration [8] shown in outline under 9% wind penetration, where green dots mark actual wind farms. We set standard deviation to be 0.3 of the mean for each wind source. Our CC-OPF (with 1% of overload set as allowable) resolved the case successfully (no overloads) and was computed in seconds, while the standard OPF showed 8 overloaded lines, all marked in color. Lines shown orange are at 4% chance of overload. There are two dark red lines which are at 50% of the overload while other (dark orange) lines show values of overload around 10%. . . . .	29
5.1	118-bus case with four wind farms (green dots; brown are generators, black are loads). Shown is the standard OPF solution against the average wind case with penetration of 5%. Standard deviations of the wind are set to 30% of the respective average cases. Lines in red exceed their limit 8% or more of the time. . . .	58
5.2	39-bus case. Red lines indicate high probability of flow exceeding the limit under the standard OPF solution. Generators are shades of blue, with darker shades indicating greater absolute difference between the chance-constrained solution and the standard solution. . . . .	60
5.3	39-bus case with four wind farms (green dots; brown are generators, black are loads). Lines in red are at the maximum of $\epsilon = .02$ chance of exceeding their limit. The three cases shown are left to right: .1%, 8%, and 30% average wind penetration. With penetration beyond 30% the problem becomes infeasible. . . .	62
5.4	30-bus case with three wind farms. The case on the left supports only up to 10% before becoming infeasible, while the one on the right is feasible up to 55% penetration. . . . .	62
5.5	Maximum probability of overload for out-of-sample tests. These are a result of Monte Carlo testing with 10,000 samples on the BPA case, solved under assuming normally distributed fluctuations and desired maximum chance of overload at 2.27%. . . . .	64

5.6	BPA case solved with average penetration at 8% and standard deviations set to 30% of mean. The maximum probability of line overload desired is 2.27%, which is achieved with 0 forecast error on the graph. Actual wind power means are then scaled according to the x-axis and maximum probability of line overload is recalculated (blue). The same is then done for standard deviations (green). . . .	64
5.7	Number of lines that are overloaded with given probability values in simulation of 2746 bus Polish power grid with 20% wind penetration distributed over 50 wind farms. Under the standard OPF, five lines are overloaded half of the time, and two lines are overloaded more than ten percent of the time, constituting a situation with unacceptable systemic risk. Under the chance-constrained OPF, the largest overload probability is <i>fifty</i> times smaller than in the case of standard OPF. Moreover, the cost increase is by less than five percent. . . . .	65
5.8	9-bus case, 25% average penetration from two wind sources. With shifting winds, the flow on the orange line changes direction with a large absolute difference. . .	67
7.1	The attack problem is non-convex and can have multiple local optima. The left image shows all attacks on lines $\{2, 3\}$ and $\{3, 4\}$ from case5 up to a maximum of $\bar{\gamma} = 2.5$ . The global optimum is $(2.5, 2.5)$ but there is a local optimum at $(.66, 0)$ . Restricting to the x-axis or the dotted line $\gamma_{34} = 2.5 - \gamma_{23}$ results in the right image, with different local and global optima; see the text for details. Random restarts of the Frank-Wolfe procedure help to find a good local optimum. . . .	80
7.2	Two different line search approaches for the Frank-Wolfe algorithm applied to minimizing $f(x, y) = (x - 2)^2 + 5(y - 2)^2$ in the region $x \geq 0, y \geq 0, x + y \leq 3$ . An exact line search suffers from zig-zagging and takes many iterations to converge; a simple back-tracking line search quickly finds the optimum. . . . .	84
8.1	case9 from MATPOWER, rendered by graphviz [34]. Generator buses are in green, load buses are in black. Buses with neither demand nor generation are grey. . . .	88

8.2	Three different attacks on case9, where a red dotted line has its impedance multiplied by 3.5. $\mathcal{F}$ is the attacker’s objective, in this case maximizing the 1-norm of phase angle differences across branches (the units are degrees). $\mathcal{H}$ is the objective for the OPF inner problem. The unattacked grid has $\mathcal{F} = 33.249$ and $\mathcal{H} = 0$ . . . . .	90
8.3	As the attack grows stronger, many different indicators of grid health grow weaker. “Null” is the unattacked grid; see the text for definitions of weak, median, and strong attacks. Branch apparent power is defined here as $S = \frac{1}{2} \left( \sqrt{P_f^2 + Q_f^2} + \sqrt{P_t^2 + Q_t^2} \right)$ – the average of the apparent power injected into the line from the ‘to’ bus and the ‘from’ bus. . . . .	91
8.4	case57 from MATPOWER with generator buses in green. Pictured are two attacks discovered with the branching procedure in section 8.2. Observe that the attacks are qualitatively different: the attack on the left creates two islands, while the attack on the left focuses on generators and heavily loaded lines. . . . .	92
8.5	The branching procedure applied to case57. The first of each pair of lines indicates which branches are excluded from the attack. The second is of the form: objective [branch_number attack_value; branch_number attack_value; . . .]. Forcing certain lines to be excluded from the attack may find other sets of vulnerable lines, and sometimes improves the strength of the attack. . . . .	93
8.6	Histogram of phase angle differences under different attacks on case118. Notice that the Frank-Wolfe attack increases the phase angle difference on many branches to over 20 degrees. . . . .	95
8.7	The IEEE 118-bus test system. (Image source: [1].) . . . . .	96
8.8	Sample output from the Frank-Wolfe attack procedure on case118 from a zero start. . . . .	96
8.9	Output from the Frank-Wolfe attack procedure on case2383wp from a zero start. Included are the step-size used in the line-search step as well as run times (in seconds) for the line search and gradient steps. . . . .	97

# List of Tables

- 2.1 Solutions to the three-bus system. . . . . 17
  
- 4.1 Performance of cutting-plane method on a number of large cases. . . . . 54
- 4.2 Typical convergence behavior of cutting-plane algorithm on a large instance. . . . . 55
  
- 5.1 Impact of penetration, 39-bus case . . . . . 59
  
- 7.1 A typical example comparing the effectiveness of a derivative-free optimization method to our gradient-based method on the attack problem. The Frank-Wolfe method finds better attacks and requires far fewer function evaluations. Runtimes were computed from a single core of an Intel Xeon X5570 2.93GHz machine. . . . . 85
  
- 8.1 Attack problem parameters for case9 example . . . . . 88
- 8.2 Changes in power generation after the strong attack.  $P_g$  is real power output,  $Q_g$  is reactive. In the strong attack, lines {1,4}, {5,6}, and {8,9} have their impedance multiplied by 3.5. . . . . 89
- 8.3 Objective values for different attacks on case118. See the text for the definitions of the different attacks. . . . . 94

# Chapter 1

## Introduction

The development of the power grid and the electricity it provides has led to tremendous economic and social progress, dramatically improving quality of life across the globe. It underlies virtually every aspect of our modern economy and society.

The energy landscape, however, is changing dramatically. One significant change is that much more of our energy is coming from renewable sources. Spurred by improvements in technology and concerns with climate change, the U.S. has increased its solar electricity generation by 20-fold since 2008, and more than tripled electricity generation from wind energy [9]. This trend is expected to continue across the world's power grids. For example Germany, a world leader in renewable energy, expects to grow its renewable penetration from 30% in 2014 to 80% in 2050.

Such high renewable penetration creates challenges for power grids like those in the U.S. which were built under the assumption of almost entirely dispatchable generation. The intermittent nature of many renewable energy sources requires the power grid to adapt from this old design to maximally take advantage of these energy sources in a safe and efficient manner.

Innovations in smart grid technology, of which increasing renewable energy plays a large part, are leading to an ever more complex power grid. At the same time, demand for power continues to increase at a faster pace than supporting infrastructure, so that the grid operates closer to capacity with smaller safety margins [42]. As a consequence, power outages and quality

disturbances are becoming more frequent, costing the U.S. economy an estimated \$80 to \$188 billion a year [12]. Recent large-scale blackouts [2] [5] provide a clear reminder of this problem. The growing complexity of the grid suggests that without new ideas and measures, these blackouts will become more frequent and more costly.

This dissertation focuses on two specific problems related to the reliability of the modern power grid. In part I we investigate the economic dispatch problem with uncertain power sources. In this, we add risk awareness to the classic problem of determining how much power generators should produce to meet demand most efficiently. We do this by explicitly modeling the uncertainty of intermittent power sources and directly addressing the risk of line flow overages when power outputs vary significantly from their forecasted values.

In part II we investigate the “the power grid attack problem”: aiming to maximize damage to the grid, how would an attacker distribute a budget of “damage” across the power lines? It is a continuous and therefore more tractable version of the “N-k” problem, which seeks to find a set of  $k$  lines in a grid with  $N$  lines whose removal will maximally disrupt the grid. By helping to identify vulnerable components in the power grid, the study of this problem can provide useful information to grid planners seeking to improve the resilience of the power grid to outages and disturbances.

# Chapter 2

## Power grid preliminaries

This chapter introduces the background material on power systems needed to understand the chapters that follow. The chapter describes basic power grid concepts, the power balance equations, and the load flow and optimal power flow problems that use them. It also includes some examples of simple power flow situations to help provide intuition as to how power flows behave.

### 2.1 Structure of power grids

The power grid can be thought of as a graph, with power lines as edges and generators and loads as vertices. The graph has two systems: 1) a high voltage (hundreds of kV) *transmission system* connecting power plants to electrical substations and 2) a medium voltage (tens of kV) *distribution system* connecting the substations to individual consumers, at which point the voltage is lowered to 120V in the U.S. The voltage levels of the different systems take advantage of different points in the trade-off between the efficiency of high voltage transmission and the safety of low voltage distribution (which occurs closer to buildings and people).

The graph of the medium voltage distribution system typically has a tree structure which can be exploited for certain problems. The problems considered in this thesis focus on the transmission system, however, which almost always includes cycles. Both systems use primarily alternating current (AC) power lines, though direct current (DC) is occasionally used in particular

situations in the transmission system.

Some terminology: the term *bus* is used to describe a vertex in the graph, usually either a generator, a load, or a transformer. The term *branch* refers to an edge in the graph – a power line.

As a matter of convenience, a rescaling of variables known as the *per unit* (or p.u.) system is almost always used in power systems analysis [17]. For example, power quantities may be scaled by 100MVA and voltages by 135KV. This is especially convenient when there are many transformers, since the per-unit quantities do not change on either side of a transformer. The system is also useful when there are many voltage levels involved; because voltage magnitudes will be close to their nominal values under stable conditions (see the next section), they are all close to 1.0 p.u. in the per-unit system, and it is easy to quickly identify buses with undesirable voltage levels.

## 2.2 Power grid stability requirements

### 2.2.1 Frequency stability

The power grid operates at a fixed frequency, 60 Hz in the U.S., with all generators producing power in sync. If two generators lose synchronism or “fall out of step,” they no longer generate voltages of the same frequency, and effective exchange of power ceases. Furthermore, because generators are specifically designed and calibrated to operate within a narrow band ( $\approx \pm 0.05\text{Hz}$ ) of the prescribed frequency, large deviations in frequency can be very damaging to the equipment. For these reasons, automatic control systems (see section 3.1 for more) act very quickly to synchronize the grid to the desired frequency. In some cases, it may be necessary to trip a generator (detach it from the grid) in order to protect it from unsafe frequency deviations, which can lead to further frequency deviations and a dangerous vicious cycle.

## 2.2.2 Voltage stability

Similarly, the grid operates most efficiently when all buses are kept at voltage magnitudes of 1.0 p.u. When grid conditions force voltages to stray too far from 1.0, a number of problems arise. This type of disturbance usually appears as a voltage drop and when voltages become low enough the system cannot maintain stability. This situation is known as *voltage collapse*. It is a well-studied phenomenon [30] involved in numerous historical blackouts, in particular the 2003 Northeast blackout [2].

## 2.2.3 Line stability

Power lines can fail when they carry too much power, most commonly in the case of thermal overload. Excessive power over a line will cause it to overheat and the conducting material to degrade. But before that occurs, the line will sag and become much more likely to touch or create an arc with a foreign object, leading to a short circuit. It takes some time for power lines to heat up, and this usually does not become dangerous for minutes or hours, leaving enough time for grid operators to shed load or redispatch power. In general, it is not a great risk when the flow on a line exceeds the limit for a short period of time. To help address the problem of thermal overload, power lines come with a standard rating and one or more emergency ratings which determine the amount of power that can be safely transmitted. Each level of rating has a maximum amount of time that the line can carry that much power before grid operators need to relieve it.

A more pressing risk to grid stability from power lines is high phase angle difference. When the phase angle difference between two buses becomes too high, the two buses will lose synchronism, leading to the problems discussed above for frequency stability. The upper limit to the phase angle difference depends on a number of factors, but it is typically more restrictive for longer lines [17].

For example, in the 2011 San Diego blackout, a long line was accidentally opened. The grid then adjusted and entered a dangerous state, while simultaneously the phase angle difference

across the line grew very large. The angle was too large for operators to reclose the line without losing synchrony and an immediate failure of the grid, so it remained open. The dangerous state of the grid persisted and led to a cascading blackout [5] which left millions in the greater San Diego area without electricity in what would become the largest power failure in California history.

## 2.3 The power balance equations

The flow of electric power on a network is governed by the laws of physics, and under stable conditions can be accurately described by the steady-state<sup>1</sup> power balance equations derived in this section. For other descriptions of this material, see [17] and [13]. First, some notation:

- $j = \sqrt{-1}$  is the imaginary unit
- $\mathcal{E}$  is the set of all lines (edges) in the grid
- $\mathcal{V}$  is the set of all buses (vertices) in the grid
- $\mathcal{G}$  is the set of generator buses
- $\mathcal{D}$  is the set of load (demand) buses
- $k$  and  $m$  are bus indices
- $K$  is the set of buses adjacent to bus  $k$ , including bus  $k$
- $\Omega_k$  is the set of buses adjacent to bus  $k$ , excluding bus  $k$
- $km$  refers to the power line from bus  $k$  to bus  $m$
- $\{k, m\}$  also refers to the power line between buses  $k$  and  $m$ , but as an unordered pair
- $\theta_{km} = \theta_k - \theta_m$  is the phase angle difference between buses  $k$  and  $m$

---

<sup>1</sup>The problems addressed in this thesis do not consider transients and other behaviors which can be described by differential equations.

- $I_{km}$  is the current injected into the line from bus  $k$  to bus  $m$
- $I_k = \sum_{m \in \Omega_k} I_{km}$  is the total current injected into the grid from bus  $k$
- $S = P + jQ$  is complex power; the real part  $P$  is called the *active power* and the imaginary part  $Q$  the *reactive power*. As with current, depending on the subscript this refers to either power along a line or the total amount injected into the grid from a bus.
- $z_{km} = r_{km} + jx_{km}$  is the impedance. The real part  $r_{km}$  is the resistance and the imaginary part  $x_{km}$  is the reactance
- $y_{km} = z_{km}^{-1} = g_{km} + jb_{km}$  is called the admittance and is sometimes more convenient to use than impedance. The real part  $g_{km} = \frac{r_{km}}{r_{km}^2 + x_{km}^2}$  is called the *conductance* and the imaginary part  $b_{km} = \frac{-x_{km}}{r_{km}^2 + x_{km}^2}$  the *susceptance*.

We start with the complex voltage at bus  $k$ :  $V_k = |V_k|e^{j\theta_k}$ . Recall Ohm's and Kirchhoff's laws:

$$I_{km} = y_{km}(V_k - V_m) \quad (\text{Ohm}) \quad (2.1)$$

$$\begin{aligned} I_k &= \sum_{m \in \Omega_k} I_{km} && (\text{Kirchoff}) \\ &= \sum_{m \in \Omega_k} y_{km}(V_k - V_m) \\ &= \left( \sum_{m \in \Omega_k} y_{km} \right) V_k - \sum_{m \in \Omega_k} y_{km} V_m \\ &= \sum_{m \in K} Y_{km} V_m \end{aligned} \quad (2.2)$$

where  $Y = G + jB$  is a weighted Laplacian known as the *admittance matrix* defined by  $Y_{km} = -y_{km}$  and  $Y_{kk} = \sum_m y_{km}$ . The real and imaginary parts of  $Y$  are the conductance and susceptance matrices, with similar definitions. The admittance matrix also includes terms for transformers and shunts, which we ignore here for simplicity of presentation. Though  $Y$  as described here is symmetric, when transformers are included that may no longer be the case.

The complex power angle is by convention the voltage angle minus the current angle, so  $S = |V||I|e^{j(\theta_V - \theta_I)} = VI^*$ , where \* indicates complex conjugation. Combining this with (2.2) gives:

$$\begin{aligned}
S_k &= V_k I_k^* \\
&= V_k \sum_{m \in K} Y_{km}^* V_m^* \\
&= |V_k| \sum_{m \in K} |V_m| (G_{km} - jB_{km}) e^{j\theta_{km}}
\end{aligned} \tag{2.3}$$

Taking real and imaginary parts gives the power balance equations:

$$P_k = |V_k| \sum_{m \in K} |V_m| (G_{km} \cos \theta_{km} + B_{km} \sin \theta_{km}) \tag{2.4}$$

$$Q_k = |V_k| \sum_{m \in K} |V_m| (G_{km} \sin \theta_{km} - B_{km} \cos \theta_{km}) \tag{2.5}$$

Eqs. (2.4)-(2.5) are written with voltages in polar coordinates and power injections in the active/reactive rectangular coordinates, which is conventional. Though rectangular coordinates are generally easier to work with, the bus voltage angle  $\theta_k$  is often a quantity of direct interest. In some cases, it can be convenient to write voltages in rectangular coordinates as well, leading to a system of quadratic equations.<sup>2</sup> The rectangular coordinate representation is used in the semidefinite relaxation solution approach discussed in section 2.6.2. If we write  $V_k = |V_k|e^{j\theta_k} = e_k + jf_k$  then the power balance equations with rectangular voltage components are:

$$P_k = e_k \sum_{m \in K} (G_{km} e_m - B_{km} f_m) + f_k \sum_{m \in K} (B_{km} e_m + G_{km} f_m) \tag{2.6}$$

$$Q_k = e_k \sum_{m \in K} (-B_{km} e_m - G_{km} f_m) + f_k \sum_{m \in K} (G_{km} e_m - B_{km} f_m) \tag{2.7}$$

which are quadratic in  $e$  and  $f$ .

---

<sup>2</sup>Thus solutions to the power flow equations form an algebraic set, falling into the domain of algebraic geometry.

## 2.4 The power flow problem

Finding a solution to the power balance equations (2.4) and (2.5) for all buses amounts to solving for the voltage magnitude and angle for each bus, using whichever of the quantities  $|V_k|, \theta_k, P_k, Q_k$  that are known as well as static grid data encoded in the  $Y$  matrix. Once we know the complex voltage at each bus, the active and reactive power injections can be immediately computed from the equations. We generally split the set of buses  $\mathcal{V}$  into two sets: 1) the set  $\mathcal{G}$  of generators, where voltage magnitude and sometimes active power injections are known and 2) the set  $\mathcal{D}$  (for demand) of loads, where active and reactive power demands are known.

For a grid with  $n$  buses of which  $m$  are loads, we have  $n$  unknown phase angles and  $m$  unknown voltage magnitudes for a total of  $n + m$  unknowns. Since active power injections are known at each bus, we have  $n$  equations of the type (2.4), and since reactive power injections are known at each load,  $m$  equations of type (2.5) for a total of  $n + m$  equations.

Conservation of power prevents us from specifying  $P_k$  at every bus, however – the sum of all power injections must equal the sum of power losses on the lines, and the line losses are unknown in advance. So we leave  $P_k$  free at one bus, called the *slack bus*, to pick up the “slack” and ensure that the sum of all power injections is equal to the sum of losses. This adds an unknown variable which would make our system of equations underdetermined. But note that in the equations, phase angles only appear as differences, so we can arbitrarily assign the phase at one bus (usually  $\theta_k = 0$ ) called the *reference bus*, thus removing one unknown. For convenience we usually let the slack and reference bus be the same at  $k = 1$ . So we indeed have  $n + m$  equations and unknowns, and the hope for a unique solution; see the next section for additional discussion.

The power balance equations are typically solved via Newton’s method. In practice, the equations are often solved after a small change has occurred in the state of the grid, and the solution from the previous state is used as the initial guess, in other words a “warm start.” If a previous solution isn’t available, a “flat start” with  $|V_k| = 1, \theta_k = 0 \forall k$  will often converge. If the grid has changed significantly from the previous known state, however, Newton’s method may fail to converge, even with a warm start. Though there has been much research and progress on

solving power flows, this problem continues to be a source of frustration in the study of power grids. Section 2.6 on the optimal power flow problem, a generalization of the simple power flow problem, further discusses this issue.

### 2.4.1 Existence and uniqueness of solutions

In general the system will not have a unique solution; we will see examples in section 2.5 of systems with zero or multiple solutions. But under normal practical conditions there is usually exactly one realistic solution: the solution where the voltage magnitudes are all close to 1.0 p.u. This is because under normal conditions, voltage magnitudes vary slowly and continuously, and grid operators are actively working to keep them near 1.0.

In general, the number of possible solutions in a grid is exponential in the number of buses [43], though the actual number of solutions is typically much smaller. The state of the art does not yet have a tractable way of reliably finding all solutions. The Newton's method grid-search method illustrated in the example in section 2.5.4 is impractical for large grids. Homotopy-based methods [59] may find all solutions, but they have complexity that grows exponentially with the size of the grid even when the number of actual solutions is small. The semidefinite relaxation approach discussed in section 2.6 is more efficient, but the relaxation is not tight in general and thus it is possible for it to find "solutions" which are not physically meaningful.

### 2.4.2 DC approximation

Though Eqs.(2.4)-(2.5) can be difficult to solve in general, under normal operating conditions they can be approximated by much simpler linear equations. The *DC approximation* takes advantage of the following approximations:

1. Resistance is dominated by reactance:  $r_{km} \approx 0$  and thus  $g_{km} \approx 0$  for all lines  $\{k, m\}$
2. Voltage magnitudes are fixed and scaled to unity:  $|V_k| \approx 1 \forall k \in \mathcal{V}$
3. Voltage angle differences are small, so that:  $\sin \theta_{km} \approx \theta_{km}$  for all lines  $\{k, m\}$

Under these assumptions, the real power balance equation (2.4) simplifies to

$$\begin{aligned}
 P_k &= |V_k| \sum_m |V_m| (G_{km} \cos \theta_{km} + B_{km} \sin \theta_{km}) \\
 &\approx \sum_m B_{km} \theta_{km}
 \end{aligned}$$

We write the DC approximation to the power balance equations as

$$B\theta = P \tag{2.8}$$

which relates the real power at each bus to the susceptance matrix and the vector of phase angle differences across each line. The name DC approximation becomes more clear when we look at the contribution of a single line to (2.8):

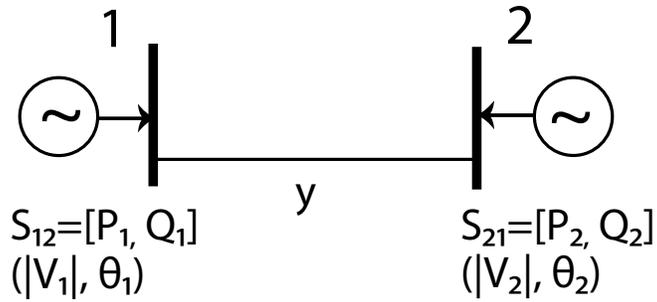
$$\begin{aligned}
 P_{km} &= B_{km} \theta_{km} \\
 &= -b_{km} (\theta_k - \theta_m) \\
 &= \frac{\theta_k - \theta_m}{x_{km}}
 \end{aligned}$$

The equation is analogous to Ohm's law applied to a resistor carrying a DC current, where in the analogy  $P$  represents the current,  $\theta$  the voltage at each terminal, and  $x$  the resistance.

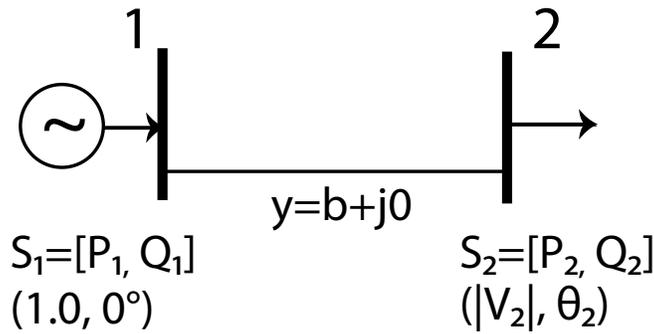
## 2.5 Power flow examples

### 2.5.1 Power circles

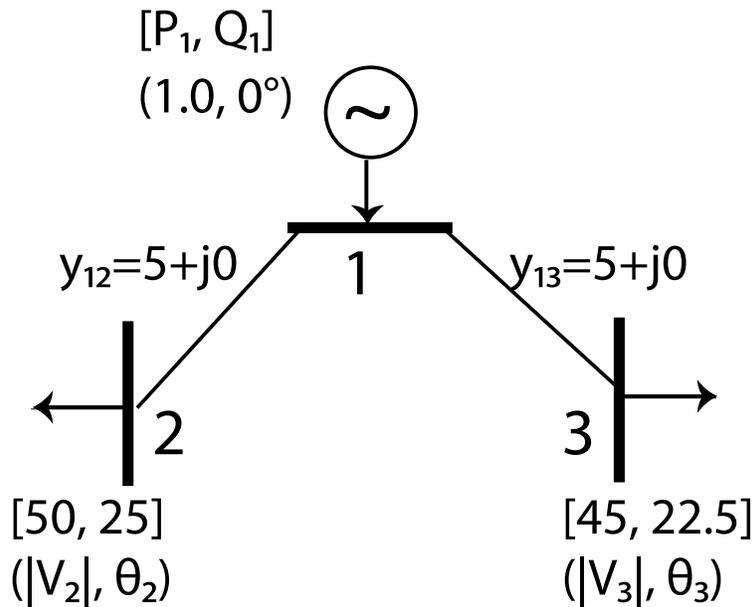
Here we introduce the concept of power circles, as described in [17]. This should provide some intuition on the behavior of power flow on a very simple grid. Consider the case of two buses connected by a single line (figure 2.1). The complex power injected into the grid by bus 1 toward



(a) Grid for the power circles example in section 2.5.1.



(b) Grid for the nose curves example in section 2.5.2.



(c) Grid for the multiple solutions example in section 2.5.4.

Figure 2.1: Two-bus and three-bus grids for the examples in section 2.5.

bus 2, from expanding out (2.3), is:

$$\begin{aligned} S_{12} &= y^* |V_1| (|V_1| - |V_2| e^{j\theta}) \\ &= y^* |V_1|^2 - y^* |V_1| |V_2| e^{j\theta} \end{aligned} \quad (2.9)$$

where we have removed the subscripts from  $y$  and  $\theta$ , so  $y = y_{12}$  and  $\theta = \theta_{12} = \theta_1 - \theta_2$ .

It is convenient to think of bus 1 as sending power to bus 2, and bus 2 as receiving power from bus 1. In this case, we can take the equivalent of (2.9) and flip the sign to get the power received by bus 2:

$$-S_{21} = -y^* |V_2|^2 + y^* |V_1| |V_2| e^{-j\theta} \quad (2.10)$$

The power loss on the line is

$$S_{12} + S_{21} = y^* |V_1 - V_2|^2 \quad (2.11)$$

Observe that with zero resistance, i.e.  $y$  is purely imaginary, there is no real power loss.

To get some understanding of how the power flow behaves on such a system, let us fix  $y$ ,  $|V_1|$ , and  $|V_2|$  – as though the two buses are both generators with a given transmission line. The power flow is then purely a function of  $\theta$ . Equations (2.9) and (2.10) are then of the form

$$S_{12} = C_1 - B e^{j\theta} \quad (2.12)$$

$$-S_{21} = C_2 + B e^{-j\theta} \quad (2.13)$$

where

$$C_1 = y^* |V_1|^2 \quad C_2 = -y^* |V_2|^2 \quad B = y^* |V_1| |V_2|$$

The plot of the circles in (2.13) is shown in figure 2.2. These  $S_{12}$  and  $-S_{21}$  circles are known as sending-end circles and receiving-end circles. The centers are  $C_1$  and  $C_2$ , and each has radius

$|B|$ . Some observations:

1. Increasing  $\theta$  from 0 increases the real power sent and received, up to a maximum.
2. In many cases (see section 2.4.2), reactance dominates resistance, and so as an approximation  $y$  is purely imaginary and the circles are vertically aligned. In this case  $P_{12} = -P_{21}$  – there is no active power loss.
3. To increase the amount of active power transmitted, we can either increase the voltage magnitudes  $|V_k|$  or increase  $|y|$ .

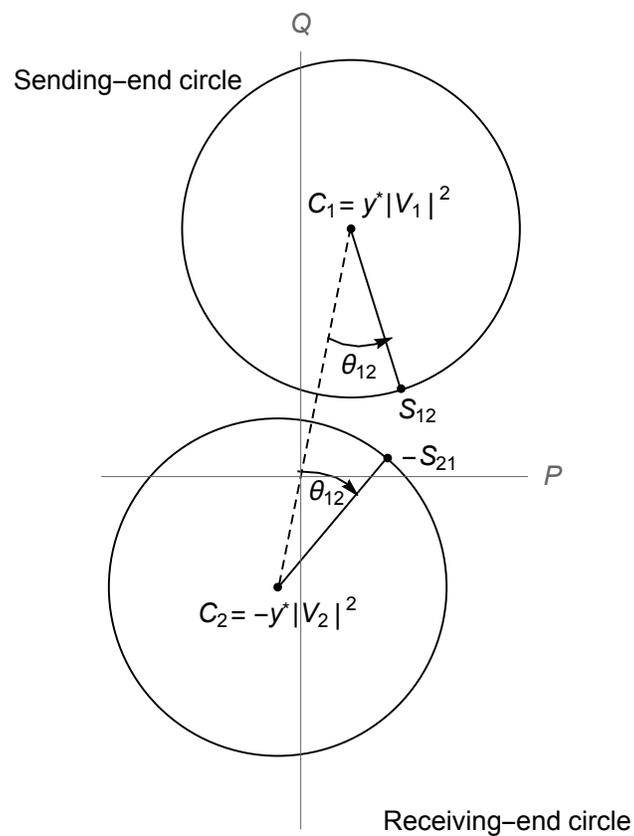


Figure 2.2: Sending-end and receiving-end circles from [17]. Both circles have radius  $|B| = |y||V_1||V_2|$ .

## 2.5.2 Nose curves

Here we introduce the concept of nose curves as described in [29]. For more on this topic in the context of voltage stability analysis, see [2, p. 35]. To start, consider a grid with two buses and a single branch (figure 2.1) with no resistance (and hence no real power loss, i.e.  $g = 0$ ). Let bus 1 be a generator with  $|V_1| = 1, \theta_1 = 0$  and let bus 2 be a load. The power flow equations at bus 2 are:

$$P_2 = bV_2 \sin \theta_2 \quad (2.14)$$

$$Q_2 = b(V_2^2 - V_2 \cos \theta_2) \quad (2.15)$$

Solving for  $V_2$  in terms of  $\theta_2$  gives

$$V_2 = \cos \theta_2 + \frac{1}{\alpha} \sin \theta_2 \quad (2.16)$$

where  $\alpha = P_2/Q_2$  is the *load factor*. If we consider different load factors, we get different contours of solutions in the  $P_2 - V_2$  plane known as *nose curves*. See figure 2.3, where we have set the susceptance to  $b = -10$ . Notice that depending on the value of  $P_2$  there are either 0, 1, or 2 solutions, and in particular that for a fixed load factor, there is a maximum amount of real power that can be delivered to the load bus beyond which the flow is infeasible.

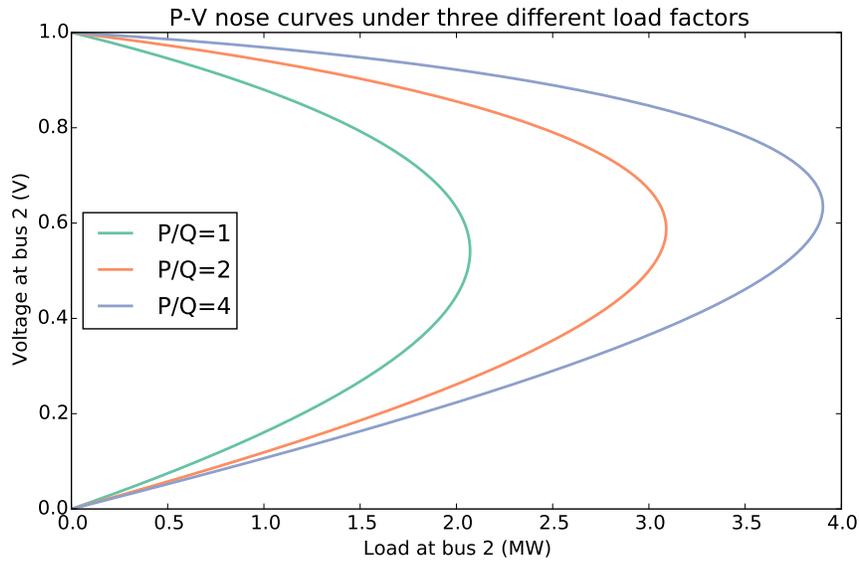


Figure 2.3: Nose curves for two bus example from [29] with  $b = -10$ . Notice how as  $P_2$  increases from zero there are first two solutions, then one solution, and finally no solutions for each load factor.

### 2.5.3 No solution

As a simple example of a case with no solution, consider the previous two bus scenario with  $P_2 = Q_2 = -b$ . In this case we have  $V_2 = \cos \theta_2 + \sin \theta_2$  from (2.16), and then from (2.15) we need a  $\theta_2$  so that  $(\sin \theta_2 + \cos \theta_2) \sin \theta_2 = -1$ . But the left-hand side has a minimum of  $(1 - \sqrt{2})/2 \approx -.207$  and so no solution exists.

### 2.5.4 Multiple solutions

In this example, also from [29], we provide an illustration of the difficulties of relying on Newton's method to find power flow solutions, even on a very simple grid. The grid we consider here adds another load bus to the grid from the previous example, for a total of three buses and two branches (see Fig. 2.1). For different initial guesses, the method may converge to one of four different solutions.<sup>3</sup> And two very close initial guesses may converge to two very different

<sup>3</sup> [29] reports that Newton's method fails to converge at all for some of the input data considered here; we found that not to be the case, though it sometimes required thirty or more iterations to converge.

solutions.

The two lines are lossless with susceptance  $b = -5$ . The loads are 50 W and 25 MVAR at bus 2 and 45 MW and 22.5 MVAR at bus 3. The four solutions are:

Solution	$ V_1 $	$\theta_1$	$ V_2 $	$\theta_2$	$ V_3 $	$\theta_3$
1	1.0	0	.944	5.9°	.946	5.7°
2	1.0	0	.082	-62.3°	.462	-14.8°
3	1.0	0	.450	-16.0°	.074	-63.1°
4	1.0	0	.120	-56.5°	.105	-58.7°

Table 2.1: Solutions to the three-bus system.

Depending on the initial values in Newton's method for  $V_2$ ,  $V_3$ ,  $\theta_2$ , and  $\theta_3$ , a different solution is found. In figure 2.4, the left image shows the solution space with initial values  $\theta_2 = \theta_3 = 0$  and the right image shows it with  $\theta_2 = -5.73^\circ$  and  $\theta_3 = -11.46^\circ$ .

The right image in figure 2.4 shows some unusual behavior, illustrating a known property of Newton's method: the edges of the basins of attraction are fractal. The upper right contains a large contiguous region (in blue) where all starting values converge to solution 1. But to the left of and below this region there are narrow bands of starting positions which instead converge to solutions 2 or 3, which are much further away. With initial angles at zero as in the left image these surprising bands do not exist. Clearly, convergence of one initial condition to a particular solution does not allow one to draw very strong conclusions regarding other nearby initial conditions.

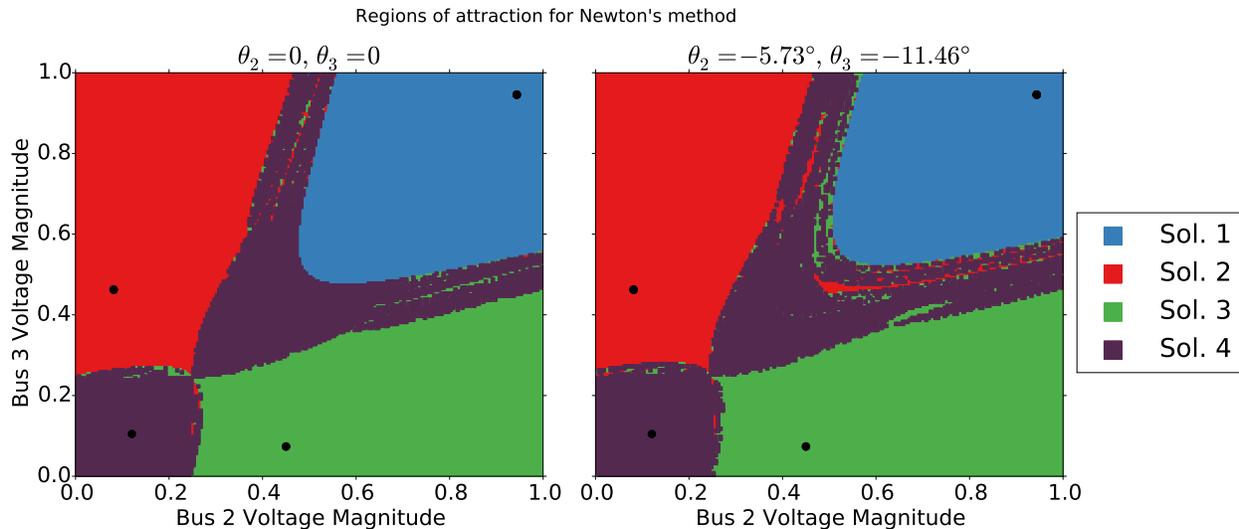


Figure 2.4: Three bus example from [29] with  $|V_1| = 0$  and  $\theta_1 = 0$ . The black dots are the four possible solutions described in table 2.1. Notice in the right image the bands of red and green near the blue region which converge to solutions 2 and 3 rather than solution 1.

## 2.6 Optimal power flow problem

The optimal power flow (OPF) problem is a generalization of the simple power flow problem in which the generator power injections and voltage magnitudes are not given. It seeks the optimal set of complex voltages at each bus and power injections at each generator bus, subject to the same power flow constraints (Eqs. (2.3)) and additional constraints on generator power injections and voltages, as well as transmission line flows. The objective is typically to minimize total operating cost though other objectives are sometimes used, such as minimizing power losses. The *economic dispatch problem*, which we examine closely in part I, is an example of the OPF that seeks to minimize cost. In the attack problem in part II, we use the OPF with a different objective to recalculate power flows after some lines have been damaged.

The OPF has been studied for several decades, with a spike in interest in recent years. It is highly non-convex and in general NP-hard. Many algorithms have been used to solve it, including gradient methods, linear programming, quadratic programming, Newton's method, augmented Lagrangian techniques, penalty methods, interior point methods, and more [23]. These methods

are generally not robust and not guaranteed to find a global optimum. Recent progress has been made with a semidefinite programming formulation [46] of the problem, discussed in section 2.6.2, which is convex and therefore in theory can guarantee either a global optimum or certificate of infeasibility. The formulation is a relaxation of the original problem, however, and the relaxation is not always tight [51], potentially giving “solutions” which do not solve the unrelaxed problem.

Using the rectangular coordinates for power injections and polar coordinates for voltages, the OPF can be formally written as follows:

$$\min_{P, Q, |V|, \theta} \sum_{k \in \mathcal{G}} c_k P_k^2 + d_k P_k \quad \text{subject to} \quad (2.17)$$

$$P_k^{min} \leq P_k \leq P_k^{max} \quad \forall k \in \mathcal{V} \quad (2.18)$$

$$Q_k^{min} \leq Q_k \leq Q_k^{max} \quad \forall k \in \mathcal{V} \quad (2.19)$$

$$\sqrt{P_{km}^2 + Q_{km}^2} \leq S_{km}^{max} \quad \forall \{k, m\} \in \mathcal{E} \quad (2.20)$$

$$V_k^{min} \leq |V_k| \leq V_k^{max} \quad \forall k \in \mathcal{V} \quad (2.21)$$

$$P_k = |V_k| \sum_{m \in K} |V_m| (G_{km} \cos \theta_{km} + B_{km} \sin \theta_{km}) \quad \forall k \in \mathcal{V} \quad (2.22)$$

$$Q_k = |V_k| \sum_{m \in K} |V_m| (G_{km} \sin \theta_{km} - B_{km} \cos \theta_{km}) \quad \forall k \in \mathcal{V} \quad (2.23)$$

Note that the last two constraints are simply the power balance equations (2.4) and (2.5). The objective (2.17) is usually a separable convex quadratic function of the power injections at the generator buses. For load buses,  $P_k$  and  $Q_k$  are given and thus not decision variables; for convenience of notation we set  $P_k^{min} = P_k^{max}$  and likewise for  $Q_k$  at each bus  $k \notin \mathcal{G}$ . Constraints (2.21)-(2.23) can alternatively be written using rectangular coordinates for voltages as in (2.6)-(2.7).

### 2.6.1 Difficulties solving the OPF

Under the non-relaxed formulation of the OPF, with difficult problems the solver may not converge. This will often be the case when the problem is in fact infeasible, but the solver will only

report that it encountered numerical difficulties or simply failed to converge. The solvers cannot guarantee that the problem is infeasible. For some applications, such as the economic dispatch problem, this may not be a huge issue: configurations of the grid where the OPF solver fails to converge can all be considered “bad” and adjusted until it does converge. For other applications (e.g. the attack problem in part II), lack of convergence is a problem because we do not know whether the configuration of the grid being considered is truly vulnerable (the OPF is infeasible – there is no possible power flow) or it is relatively harmless and the solver is just having trouble. Another potential problem with the standard formulation of the OPF is that a solution found via a local method may fail to be a global optimum. [22] includes many examples of OPF problems with suboptimal local solutions.

Significant progress in the robustness of OPF solvers has been made recently, in particular the semidefinite relaxation discussed in the next section. These new techniques have their own drawbacks, however, and software implementations are not as mature. For this dissertation, the OPF is an important building block of the attack problem in part II. We used the interior-point solver IPOPT [66] which performs well in practice. For some small problems, we used the semidefinite programming approach to obtain certificates of infeasibility. We use the OPF with the DC approximation in part I for the economic dispatch problem; because the DC assumptions make the problem linear, the difficulties described here do not arise. Section 2.7 describes some specific examples of difficult OPF problems.

## 2.6.2 Semidefinite relaxation of the OPF

A recent development that has renewed interest in solution methods for the OPF is the semidefinite relaxation. It is a convex formulation of the problem, and thus has the potential to be much more robust than conventional solution methods. In this section we introduce semidefinite programming, and very briefly describe the semidefinite relaxation of the OPF as well as its shortcomings. For a more thorough treatment, see e.g. [46–48, 51].

Semidefinite programming (SDP) is a method that optimizes a linear objective function over

the intersection of a cone of positive semidefinite matrices and an affine plane. SDP problems are convex and can be efficiently solved with interior point methods. Some free software that can solve SDPs include SeDuMi [63] and SDPT3 [64].

Denote by  $S^n$  the set of symmetric  $n \times n$  matrices, and for two  $n \times n$  matrices  $A$  and  $B$ , the Frobenius inner product:

$$A \bullet B = \text{trace}(AB) = \sum_{i=1}^n \sum_{j=1}^n A_{ij} B_{ij}$$

A general semidefinite program can be written as

$$\min_{W \in S^n} C \bullet W \tag{2.24}$$

$$\text{subject to } A_i \bullet W = b_i, \quad i = 1, \dots, m \tag{2.25}$$

$$W \succeq 0 \tag{2.26}$$

where  $W$  is the matrix decision variable,  $C \in S^n$  is a given cost matrix,  $A_i \in S^n$  are given constraint matrices with the constraint vector  $b \in \mathbb{R}^m$ , and  $\succeq$  indicates that a matrix is positive semidefinite.

We now develop the semidefinite relaxation of the OPF, using notation from [18]. Let  $v \in \mathbb{R}^{2n}$  (where  $n$  is the number of buses) be the vector  $(e_1, \dots, e_n, f_1, \dots, f_n)^T$  – i.e. the vector of rectangular voltage coordinates. Observe that because the rectangular power balance equations (2.6)-(2.7) are quadratic in  $e$  and  $f$  they can be written as

$$P_k = v^T M_k v \tag{2.27}$$

$$Q_k = v^T N_k v \tag{2.28}$$

for suitable matrices  $M_k$  and  $N_k$ , which simply combine the appropriate elements of the admit-

tance matrix  $Y = G + jB$ . Then the OPF objective (2.17) can be written

$$\min \sum_{k=1}^n c_k (v^T M_k v)^2 + d_k (v^T M_k v)$$

where  $c_k = d_k = 0$  whenever  $k \notin \mathcal{G}$ . We can rewrite this as

$$\begin{aligned} & \min \sum_{k=1}^n \kappa_k \\ \text{s.t. } & c_k (v^T M_k v)^2 + d_k (v^T M_k v) \leq \kappa_k, \quad k = 1, \dots, n \end{aligned} \quad (2.29)$$

It can be shown that an equivalent matrix form of the constraint 2.29 is:

$$\begin{pmatrix} \kappa_k - d_k v^T M_k v & -\sqrt{c_k} v^T M_k v \\ -\sqrt{c_k} v^T M_k v & 1 \end{pmatrix} \succeq 0 \quad (2.30)$$

Let  $W$  be the  $\mathbb{R}^{n \times n}$  matrix decision variable. The idea is that  $W = vv^T$ , but this will not always be the case as we will discuss below. For simplicity, we assume that all generators have zero load. The semidefinite relaxation of the OPF without line flow constraints can then be written as follows:

$$\begin{aligned} & \min_W \sum_{k=1}^n \kappa_k \\ \text{s.t. } & W \succeq 0, \text{ and for } k = 1, \dots, n : \end{aligned} \quad (2.31)$$

$$\begin{pmatrix} \kappa_k - d_k M_k \bullet W & -\sqrt{c_k} M_k \bullet W \\ -\sqrt{c_k} M_k \bullet W & 1 \end{pmatrix} \succeq 0, \quad (2.32)$$

$$P_k^{min} \leq M_k \bullet W \leq P_k^{max}, \quad (2.33)$$

$$Q_k^{min} \leq N_k \bullet W \leq Q_k^{max}, \quad (2.34)$$

$$(V_k^{min})^2 \leq T_k \bullet W \leq (V_k^{max})^2 \quad (2.35)$$

where  $T_k$  is a diagonal matrix with ones in position  $k$  and  $k + n$  and otherwise zero, so that  $T_k \bullet vv^T = e_k^2 + f_k^2$ . Including the line flow constraints requires a construction similar to that of the other constraints.

If we replace the constraint  $W \succeq 0$  with the rank-one constraint

$$W = vv^T \tag{2.36}$$

the formulation is exact and not a relaxation. Constraint 2.36 is not convex, however, and thus cannot be enforced by an SDP solver. The semidefinite relaxation is tight if the global optimum  $W$  happens to have rank one. This is the case for most practical instances of the problem, and under a number of appropriate sufficient conditions [47, 48]. In general, however, the relaxation can produce an optimum  $W$  with rank greater than one, and thus fail to produce a meaningful solution  $v$ . Furthermore, the SDP relaxation can be feasible when the original OPF is infeasible [44]. [51] and [22] include additional practical examples where the relaxation fails because  $\text{rank}(W) > 1$ .

Another significant drawback of the semidefinite approach is that it is much more computationally expensive: the number of decision variables increases from order  $n$  to order  $n^2$ . In practice, for large and complex grids the SDP formulation can be prohibitively slow to solve for some applications. Progress continues to be made, however, and [45] describes other relaxations of the OPF which are much faster to solve.

## 2.7 Difficult optimal power flow examples

To reinforce the difficulty of the OPF, this section will examine some problems which prove troublesome for numerical solvers. We will provide examples for the following situations:

1. The solver converges to a non-optimal local solution
2. Two different solvers converge to two different solutions

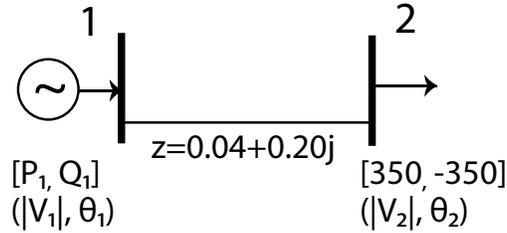


Figure 2.5: Two bus grid from [22]. Note that we use the impedance  $z$  rather than the admittance  $y$  in this example.

3. The solver fails to converge after a very large number of iterations

### 2.7.1 Two-bus OPF with local solutions

Consider the two bus grid in Fig. 2.5 from [22]. There is a single generator at bus 1 and the objective is to minimize the active power output. Voltage magnitude limits on each bus are set to  $[.95, 1.05]$  and there are no line flow limits. This very simple OPF has a feasible region consisting of two disconnected segments in the  $|V_1| - |V_2|$  plane. Setting the initial guess for a standard OPF solver in one region results in the global optimum; starting in the other region leads to a more costly local optimum. See Fig. 2.6.

Furthermore, as reported in [22], setting the voltage magnitude limit for bus 2 to a value in the range  $[.977, 1.034]$  causes the SDP solver to fail. Conventional local solvers correctly converge to the only remaining feasible optimum at  $(.95, .976)$  in this case.

### 2.7.2 Different solvers converge to different solutions

In the modification described in [22] to the nine bus grid case9 from MATPOWER [72], they raise the reactive power generation bounds from  $-300$  MVAR to  $-5$  MVAR and scale all loads by 60%. This gives rise to four local solutions. Solving this case with the trust region solver TRALM from [68] results in the globally optimal solution. Solving it instead with the interior point method PDIPM, also from [68], results in the worst of the four local solutions, which has a cost 37% greater than the global optimum.

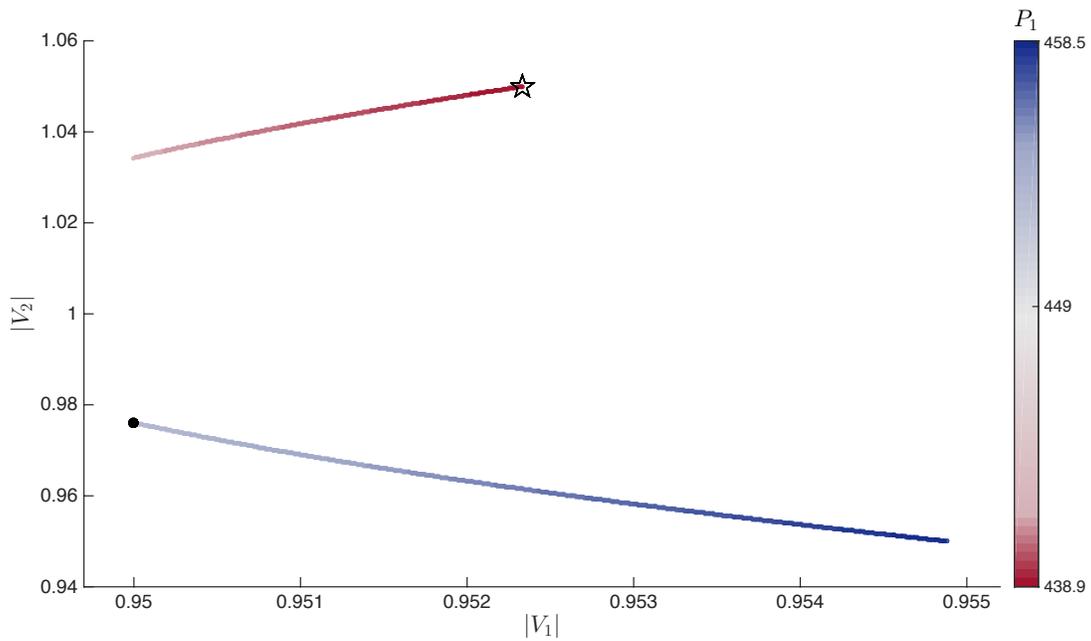


Figure 2.6: Two bus example from [22]. Note that the set of feasible points  $(V_1, V_2)$  is disconnected. Searching locally along the blue feasible region yields the non-optimal local solution at  $(.95, .976)$  marked with the dot. Searching in the red feasible region gives the global minimum marked by the star at  $(.952, 1.05)$ .

### 2.7.3 Failure to converge after many iterations

This particular example comes from our investigation of the attack problem in part II. The 57-bus grid in case57 from MATPOWER is modified so that the lines 35, 36, and 46 have their impedance multiplied by 4. Line flow constraints are removed and voltage magnitude constraints at load buses are relaxed to  $[-.5, 1.5]$ . Voltage constraints at generators are set to  $\pm 1\%$  of their normal values.

At iteration 8017 and after over a minute of computation, IPOPT reports that it believes the problem is infeasible. This required several iterations of increasing the maximum number of iterations because the default maximum is set to 250, which works well for the vast majority of problems. In contrast, the SDP solver reports the problem is infeasible in four seconds and 31 iterations. In the attack problem, small variations of this OPF are computed thousands of times as a subroutine, usually converging in a fraction of a second. The performance of IPOPT for this problem illustrates the difficulty of using the OPF as a subroutine with the current technology.

# Part I

## The economic dispatch problem with uncertain power sources<sup>4</sup>

---

<sup>4</sup>Adapted from *Chance-Constrained Optimal Power Flow: Risk-Aware Network Control under Uncertainty* by Daniel Bienstock, Michael Chertkov, and Sean Harnett. SIAM Review 56 (2014) [19].

# Chapter 3

## Introduction

The smart grid – the use of computer-based control and automation technologies to manage the increasing complexity of the modern power grid – is well on its way. A large motivation for the smart grid and a major contributor to the increased complexity is our desire for greater use of renewable energy. But the large-scale introduction of renewables brings with it the risk of large, random variability. This is a condition that the current grid, built in the 1890s under much simpler electricity needs, was not developed to accommodate. Automatic grid control and regulations achieve remarkable robustness of operation under normal fluctuations, in particular under the forecasted inter-day demands. But as the size of power fluctuations grows, which it will with increased renewable penetration, the automatic systems can fail and must rely on human input to restore stability to the grid.

The *economic dispatch problem* (see section 2.6) is one example where greatly increased renewable penetration creates challenges for the way the grid is currently operated. Economic dispatch is used to set generator outputs over approximately 15-minute time windows so as to meet demand at minimum cost, subject to transmission and operational constraints. Estimates of the expected loads for the upcoming time window are employed in this computation. In real time, the generator outputs computed by economic dispatch are modulated by the Automatic Generation Control (AGC) system (see section 3.1) to account for variations in demand from

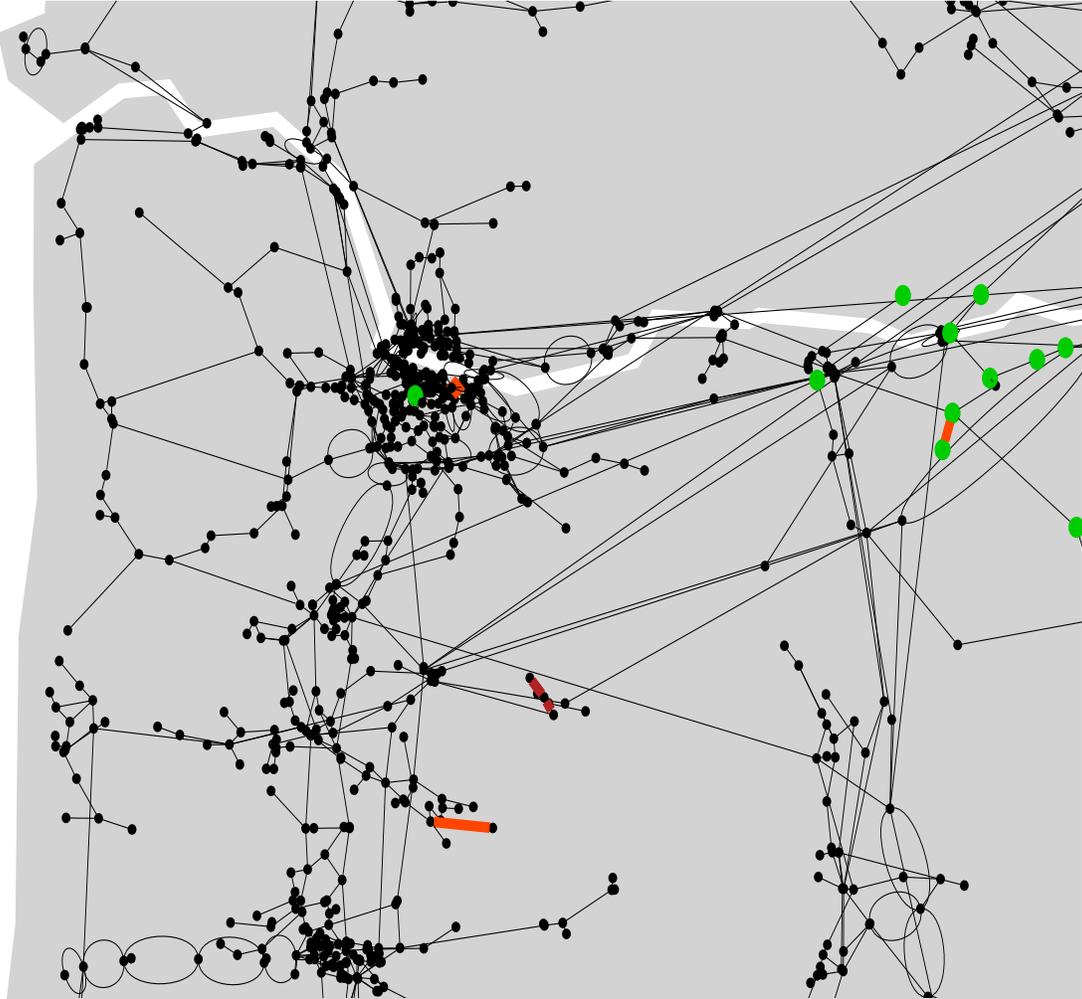


Figure 3.1: Bonneville Power Administration [8] shown in outline under 9% wind penetration, where green dots mark actual wind farms. We set standard deviation to be 0.3 of the mean for each wind source. Our CC-OPF (with 1% of overload set as allowable) resolved the case successfully (no overloads) and was computed in seconds, while the standard OPF showed 8 overloaded lines, all marked in color. Lines shown orange are at 4% chance of overload. There are two dark red lines which are at 50% of the overload while other (dark orange) lines show values of overload around 10%.

the estimates. Because AGC does not respect transmission constraints, this scheme can fail dramatically when renewables are part of the generation mix and fluctuations in their output become large. Specifically, combinations of generator and renewable outputs can conspire to produce power flows that significantly exceed transmission line ratings, increasing the chances that those lines trip (see section 2.1). If several key lines trip, the grid is likely to become unstable and experience a cascading failure, with large losses in served load. To prevent line tripping an additional scheme is employed as a part of the current operational routine. It requires a human operator in the loop and is based on direct line flow measurements: after receiving a warning, the operator may initiate an emergency action, possibly disconnecting the overheated line.

The U.S. is committed to increasing the proportion of power delivered from renewable sources [4] while safety margins (between typical power flows and line ratings) continue to shrink [42]. Without significant changes, line overloads will become much more frequent, and so the current operational paradigm is most likely unsustainable. A possible failure scenario is illustrated in figure 3.1 with U.S. Pacific Northwest regional grid data (2866 lines, 2209 buses, 176 generators and 18 wind sources), where lines highlighted in red are jeopardized (flow becomes too high) with unacceptably high probability by fluctuating wind resources positioned along the Columbia river basin.

This part of the dissertation investigates a solution by explicitly modeling fluctuating power sources in the economic dispatch problem. The non-deterministic behavior of wind make it natural to cast the problem in terms of stochastic optimization. Recent works [16,26,27] suggest that focusing on the most-likely dangerous events provides a practicable route to risk control and assessment. The approach described here goes one step further, by implicitly discovering probable realizations of line overloads and correcting them in a single step.

The particular type of stochastic optimization used in our approach is known as “chance-constrained” optimization [54], which allows for constraints stating that the probability of a certain random event is kept smaller than a target value. Our formulation of the problem minimizes the average cost of generation over the random power injections (from renewables), while specifying

a mechanism by which the dispatchable generators adjust in real-time to compensate for the power fluctuations, and guaranteeing a low probability that any line will exceed its rating. This last constraint is naturally formulated as a chance constraint. The economic dispatch problem is a particular type of optimal power flow problem (section 2.6) and though we examine only economic dispatch, our method applies more generally to OPF problems. Thus we term our approach Chance-Constrained Optimal Power Flow, or CC-OPF.

Part I of this dissertation is organized as follows. The remainder of this chapter motivates and presents the various mathematical models used to describe how the grid operates, as well as the proposed methodology. Chapter 4 explains how to solve the models. Chapter 5 then presents a number of examples to demonstrate the speed and usefulness of our approach.

## 3.1 Transmission system controls

Transmission systems (see 2.1 for more background information) balance load and generation using a complex strategy that spans three different time scales (see e.g. [17]). An essential stability requirement is that all generators operate at the same frequency – 60Hz in the U.S. Failure to maintain a common frequency leads to a loss of synchrony, and can force generators to shut down for protective reasons, potentially leading to a swift collapse of the grid.

In real time, changes in loads are observed at generators by an opposite change in frequency. Consider the case of a sudden load increase. In that case generator frequency will start to drop. The so-called primary frequency control will react to stop frequency drift by having each participating generator inject more power into the grid, proportionally to the frequency change. This reaction is swift and local, leading to stabilization of frequency across the system, however not necessarily at the nominal 60Hz value.

The task of the secondary, or Automatic Generation Control (AGC), is to undertake the adjustment of generation levels to return frequency to 60Hz. The economic dispatch algorithm typically runs as frequently as every 15 minutes providing information to the AGC, which ultimately

undertakes the adjustment of generation levels to achieve optimal (or close to optimal) control. The economic dispatch time-window thus represents the shortest time scale where actual off-line and network wide optimal computations are employed.

## 3.2 Standard generation dispatch

This section reviews and further describes the way power is usually dispatched, by means of a conventional optimal power flow. We'll refer to this scheme as simply "standard OPF". Note: to stay consistent with [19] we will use slightly different notation from the discussion in section 2.3: we'll use  $f_{km}$  rather than  $P_{km}$  for the power flow along line  $\{k, m\}$ , and  $\beta_{km}$  for susceptance rather than  $b_{km}$  so that  $f_{km} = \beta_{km}(\theta_k - \theta_m)$ .

The optimal power flow problem is introduced in section 2.6. We use the DC approximation (section 2.4.2), so we ignore reactive power and voltage magnitude. We can informally describe the DC-OPF problem as follows:

- The goal is to determine the vector of active power generated and voltage angles so as to minimize a convex quadratic objective function of the generator power outputs
- There are three types of constraints: power flow, line limit, and generation bound constraints.

Let  $d_k$  be demand (possibly zero) at bus  $k$  and  $p_k$  the power generated (also possibly zero). Then the power injected into the grid from bus  $k$  is  $p_k - d_k$ . With this notation, the DC power flow equation in matrix form is

$$B\theta = p - d \tag{3.1}$$

where recall that the susceptance matrix  $B$  is a weighted-Laplacian defined by  $B_{km} = -\beta_{km}$  and  $B_{kk} = \sum_{m \in \Omega_k} \beta_{km}$ .

For future reference, we state some well known properties of Laplacians and the power flow system (3.1).

**Lemma 3.2.1.** The following statements hold for the system (3.1):

- the sum of rows of  $B$  is zero
- if the underlying graph is connected, the rank of  $B$  is  $n - 1$
- if (3.1) is feasible, then for any index  $k$  there is a solution with  $\theta_k = 0$
- the system (3.1) is feasible in  $\theta$  if and only if total generation equals total demand

More formally, the standard DC-OPF problem can be stated as the following constrained optimization problem:

$$\text{DC-OPF: } \min_{p, \theta} \quad c(p), \quad \text{s.t.} \quad (3.2)$$

$$B\theta = p - d, \quad (3.3)$$

$$\forall k \in \mathcal{G} : \quad p_k^{\min} \leq p_k \leq p_k^{\max}, \quad (3.4)$$

$$\forall \{k, m\} \in \mathcal{E} : \quad |\beta_{km}(\theta_k - \theta_m)| \leq f_{km}^{\max}, \quad (3.5)$$

Note that the  $p_k^{\min}, p_k^{\max}$  quantities can be used to enforce the convention  $p_k = 0$  for each  $k \notin \mathcal{G}$ ; if  $k \in \mathcal{G}$  then  $p_k^{\min}, p_k^{\max}$  are lower and upper generation bounds which are generator-specific. Constraint (3.5) is the line limit constraint for  $\{k, m\}$ ;  $f_{km}^{\max}$  represents the line limit which is assumed to be strictly enforced. This conservative condition will be relaxed in what follows.

Problem (3.2) is a convex quadratic program, easily solved using modern optimization tools. The vector  $d$  of demands is *fixed* in this problem and is obtained through estimation. In practice, however, demand will fluctuate around  $d$ ; generators then respond by adjusting their output (from the OPF-computed quantities) proportionally to the overall fluctuation as discussed above in section 3.1. This scheme works well in current practice, since demands usually do not fluctuate significantly over the approximately 15-minute time scale between OPF calculations.

### 3.3 Adjusting to power fluctuations

When fluctuations in power injections grow large, the scheme in the previous section can lead to serious line overloads. This section examines more closely the details of how this happens, by describing how generator output is modulated in real time to respond to changes in demand.

Suppose we have computed, using the standard OPF, the output  $p_k$  for each generator  $k$  assuming constant demands  $d$ . Let  $\hat{d}(t)$  be the vector of real-time demands at time  $t$ , and similarly  $\hat{p}(t)$  the vector of real-time generator outputs. Then so-called “frequency control,” or more properly, the combination of primary and secondary controls (section 3.1) will achieve (on the scale of minutes) the following real-time generator outputs:

$$\hat{p}_k(t) = p_k - \rho_k \sum_m (d_m - \hat{d}_m(t)) \quad \forall k \in \mathcal{G} \quad (3.6)$$

In this equation, the quantities  $\rho_k \geq 0$  are fixed and satisfy  $\sum_k \rho_k = 1$ .  $\rho_k$  represents the proportional adjustment made by generator  $k$  to the mismatch between the forecasted and actual demand. Taking the sum of (3.6) over all  $k$ , and assuming the system is feasible so that  $\sum_k p_k = \sum_k d_k$ , we obtain

$$\sum_k \hat{p}_k(t) = \sum_k p_k - \sum_m (d_m - \hat{d}_m(t)) = \sum_k \hat{d}_k(t),$$

and thus the demand matches supply over time. The quantities  $\rho_k$  are generator dependent but essentially chosen far in advance and without regard to short-term demand forecasts.

Generator outputs are set in this hierarchical fashion: the OPF computes a base level, and then real-time adjustments are made according to (3.6) which is risk-unaware. This scheme has worked in the past because of the slow time scales of change in uncontrolled resources (mainly loads). That is to say, frequency control and load changes are well-separated. An error in the forecast of  $d$  for the next – e.g., 15 minute – period may lead to an operational problem<sup>1</sup>. This is

---

<sup>1</sup>See for example the discussions in [3, 49]

because even though the vector  $\hat{p}(t)$  suffices to meet average demands, the  $\hat{\theta}(t)$  computed from

$$B\hat{\theta}(t) = \hat{p}(t) - \hat{d}(t)$$

may give rise to real-time power flows

$$\hat{f}_{km}(t) \doteq \beta_{km}[\hat{\theta}_k(t) - \hat{\theta}_m(t)]$$

that violate the line flow constraints (3.5). Even the generator constraints (3.4) may fail to hold. This has not been considered a handicap since any resulting line trips are rare, primarily because the deviations  $\hat{d}_k(t) - d_k$  will be small in the time scale of interest. In effect, the risk-unaware approach that assumes constant demands has worked well.

Now let us consider the case where demands are constant, but some generators have fluctuating output.<sup>2</sup> Let a subset  $\mathcal{W}$  of the buses hold uncertain power sources (wind farms); for each  $k \in \mathcal{W}$ , write the amount of power generated by source  $k$  at time  $t$  as  $\mu_k + \omega_k(t)$ , where  $\mu_k$  is the forecast output of farm  $k$  in the time period of interest and  $\omega_k(t)$  is the fluctuation from the forecast over time. For ease of exposition, we will assume in what follows that  $\mathcal{G}$  refers to the set of buses holding *controllable* generators, i.e.  $\mathcal{G} \cap \mathcal{W} = \emptyset$ . Renewable generation can be incorporated into the OPF formulation (3.2)-(3.5) by simply setting  $p_k = \mu_k$  for each  $k \in \mathcal{W}$ . With constant demands and fluctuating power sources, the application of frequency control yields the following analog to (3.6):

$$\hat{p}_k(t) = p_k - \rho_k \sum_{m \in \mathcal{W}} \omega_m(t) \quad \forall k \in \mathcal{G} \quad (3.7)$$

For instance, if  $\sum_{m \in \mathcal{W}} \omega_m(t) > 0$ , that is, there is a net increase in wind output, then (controllable) generator output will proportionally decrease.

Eq. (3.7) describes how generation will adjust to wind changes under current power engineering

---

<sup>2</sup>This is essentially a matter of notation, but we make the distinction because we will use the fluctuating generator formulation in what follows.

practice. The hazard embodied in this relationship is that the quantities  $\omega_m(t)$  can be large, resulting in changes in power flows significant enough to overload power lines. The risk of such overloads can be expected to increase [42]; this is due to a projected increase of renewable penetration in the future [4], accompanied by the decreasing gap between normal operation and limits set by line capacities. Tightening the line flow limits can succeed in deterministically preventing overloads, but it also forces excessively conservative choices of the generation re-dispatch, with the potential risk of greatly increased cost and extreme volatility of the electricity markets. See e.g. the discussion in [67] on abnormal price fluctuations in markets that are heavily reliant on renewables.

### 3.4 Using chance constraints

Section 2.1 describes the nature of a line trip due to thermal overload – the result of carrying too much power for too long. The failure does not necessarily happen instantly and is difficult to model<sup>3</sup> but a trip is more likely to occur the longer the line stays overheated.

So as an enhancement to the standard OPF which assumes static demand and power generation, we would ideally address the problem in the previous section by updating the static line constraints (3.5) with constraints of the form “*the fraction of the time that the line exceeds its limit within a certain time window is small.*” Direct implementation of this constraint would require resolving dynamics of the grid over the generator dispatch time window of interest. Instead we propose the following static proxy of this ideal model – a *chance constraint*: we will require that the probability that a given line exceeds its limit is small.

To formalize this notion, we assume:

1. For each  $k \in \mathcal{W}$ , the (stochastic) amount of power generated by source  $k$  is of the form  $\mu_k + \omega_k$ , where
2.  $\mu_k$  is constant, assumed known from the forecast, and  $\omega_k$  is a zero mean independent

---

<sup>3</sup>See [2] for discussions of line tripping during the 2003 Northeast U.S.-Canada blackout.

random variable with known standard deviation  $\sigma_k$ .

Here and in what follows, we use bold face to indicate uncertain quantities. For line  $\{k, m\}$ , let a tolerance parameter  $\epsilon_{km} > 0$  be given and recall that  $\mathbf{f}_{km}$  is the (uncertain) flow. The chance constraint for line  $\{k, m\}$  is then:

$$P(\mathbf{f}_{km} > f_{km}^{max}) < \epsilon_{km} \quad \text{and} \quad P(\mathbf{f}_{km} < -f_{km}^{max}) < \epsilon_{km} \quad (3.8)$$

Likewise<sup>4</sup>, for a generator  $g$  we will require that

$$P(\mathbf{p}_g > p_g^{max}) < \epsilon_g \quad \text{and} \quad P(\mathbf{p}_g < p_g^{min}) < \epsilon_g. \quad (3.9)$$

The parameter  $\epsilon_g$  will be chosen *extremely* small, so that for all practical purposes all generator outputs will be guaranteed to stay within respective bounds.

Chance constraints [25, 50, 58] are but one possible methodology for handling uncertain data in optimization. Broadly speaking, this methodology fits within the general field of stochastic optimization. Constraint (3.8) can be viewed as a “value-at-risk” statement; the closely-related “conditional value at risk” concept provides a (convex) alternative, which roughly stated constrains the expected overload of a line to remain small, *conditional* on there being an overload (see [54] for definitions and details).

One alternative model would impose the much stronger constraint

$$P(\exists \text{ line } \{k, m\} \text{ s.t. } |\mathbf{f}_{km}| > f_{km}^{max}) < \epsilon, \quad (3.10)$$

with a single global tolerance  $\epsilon$ . Nemirovski and Shapiro [54] develop a general framework for constructing, under appropriate assumptions, a convex optimization problem with an approximate version of constraint (3.10).

---

<sup>4</sup>One could alternatively use the more conservative constraint  $P(|\mathbf{f}_{km}| > f_{km}^{max}) < \epsilon_{km}$ . This implies (3.8), and if (3.8) holds then  $P(|\mathbf{f}_{km}| > f_{km}^{max}) < 2\epsilon_{km}$ . However, (3.8) proves more tractable, and moreover we are interested in the regime where  $\epsilon_{km}$  is fairly small; thus we estimate that there is small practical difference between the two constraints; this will be verified by our numerical experiments.

Reference [70] considers the standard OPF problem under stochastic demands, using chance constraints to guarantee high probability that the system operates within acceptable bounds. The problem is tackled using a simulation-based local optimization system, with experiments using a 5-bus and a 30-bus example.

Another related study [65] describes a scenario-based system for reserve scheduling with fluctuating wind generation, using chance constraints to limit line or generator overloads. This optimization is tackled via transformation to a convex problem and a heuristic scheme, with no convergence to global optimum of the nonlinear problem guaranteed.

Chance constrained optimization has also been discussed recently in relation to the unit commitment problem, which is used to plan the operation of large generation units on the scale of hours-to-months, so as to account for long-term wind-farm generation uncertainty [56, 69, 71].

### 3.5 Uncertain power sources

In what follows, we assume that the outputs of the wind power sources vary between OPF calculations as (1) independent (2) normally distributed random variables with (3) known mean and variance. Let us examine the components of this assumption.

We assume that reasonably accurate, standard forecasts of wind speed for the next OPF time window are available at a coarse-grained level of minutes and kilometers. Along with known “power curves” for the wind turbines, which describe the relationship between wind speed and power output, this provides a forecast of the power generated by each wind farm for the next time period. This sort of forecast data is already widely used, so this is not a large assumption. The possibility of forecast errors can be handled efficiently with the so-called ambiguous chance constrained approach, which we describe in detail in [19]. This extends the formulation here, where the forecasts are constant, by letting them vary within a fixed set. Furthermore, the out-of-sample numerical simulations in section 5.5 suggest that the degradation in the chance constraints are small when data errors are small, so that the solution from our approach is not especially sensitive

to errors in the wind forecasts.

The independence of wind power fluctuations<sup>5</sup> at different sites can be justified by the fact that the wind farms are sufficiently far away from each other. For the typical economic dispatch time span of 15 minutes and typical wind speed of 10m/s, fluctuations at farms more than 10km apart are not correlated.

The assumption that the power output of a wind farm over a 15 minute period varies according to a normal distribution is less justified. We use it as a first approximation for the purpose of simplicity and because the resulting optimization problem is particularly efficient. Though once again, numerical experiments (section 5.5) suggest that the errors due to non-normal wind fluctuations are relatively small. The distribution of short-term (e.g. ten minute) average wind speed observations over long time periods is well-researched [52], and the Weibull distribution is commonly proposed as a good fit. In other words, there is a lot of data available for the distribution of  $\mu_k$  over a long period of OPF time windows. But our problem is concerned rather with the fluctuations of wind speed from the forecasted value –  $\omega_k$  – on even shorter time scales. Furthermore, we are concerned with wind power rather than speed, so turbine power curves must be taken into account. Techniques exist<sup>6</sup> to enable distributions other than the normal distribution in our methodology, but at the cost of increased complexity. Future work on this problem should revisit this assumption of normally distributed wind power fluctuations.

## 3.6 Affine Control

Since the power injections at each bus are fluctuating, we need a control scheme to ensure that generation is equal to demand at all times within the time window of interest. We assume that all governors involved in the controls respond to fluctuations in the generalized load (actual demand which is assumed frozen minus stochastic wind resources) in a proportional way, however

---

<sup>5</sup>We assume that the *fluctuations*  $\omega_k$  are independent. Of course the overall wind levels in a region can be correlated, which will be reflected in the forecasts of the average wind power  $\mu_k$ .

<sup>6</sup>Chebyshev's inequality is a rather crude technique, and there are more advanced methods from robust optimization [54].

with possibly different proportionality coefficients. Thus, we term the joint result of the primary frequency control and secondary frequency control the *affine* control. The stochastic version of Eq. (3.7) thus becomes

$$\mathbf{p}_k = \bar{p}_k - \alpha_k \sum_{m \in \mathcal{W}} \boldsymbol{\omega}_m \quad \forall k \in \mathcal{G} : \quad (3.11)$$

Here the quantities  $\bar{p}_k \geq 0$  and  $\alpha_k \geq 0$  are design variables satisfying (among other constraints)  $\sum_{k \in \mathcal{G}} \alpha_k = 1$ . Thus the generator output  $p_k$  combines a fixed term  $\bar{p}_k$  and a term which varies with wind,  $-\alpha_k \sum_{m \in \mathcal{W}} \boldsymbol{\omega}_m$ . Observe that  $\sum_k p_k = \sum_k \bar{p}_k - \sum_m \boldsymbol{\omega}_m$ , that is, the total power generated equals the average production of the generators minus any additional wind power above the average case.

Allowing  $\alpha$  to change does not reflect current practice. We allow this degree of freedom because (a) control of  $\alpha$  is available, in principle; (b) we do not set any  $\alpha_k$  to a standard (fixed) value, but instead leave it to the optimization to decide the optimal value; and (c) the freedom is optional in that our formulation allows for linear constraints on  $\alpha$ , so that it can be fixed if desired. (In some cases it may even be advantageous to allow negative  $\alpha_k$  but we decided not to consider such a drastic change of current policy in this study.)

This affine control scheme creates the possibility of requiring a generator to produce power beyond its limits. With unbounded wind, this possibility is inevitable, though we can restrict it to occur with arbitrarily small probability, which we will do with additional chance constraints for all controllable generators,  $\forall g \in \mathcal{G}$ ,

$$P(p_g^{min} \leq \bar{p}_g - \alpha_g \sum_{k \in \mathcal{W}} \boldsymbol{\omega}_k \leq p_g^{max}) > 1 - \epsilon_g. \quad (3.12)$$

### 3.7 Overview of the CC-OPF solution methodology

Our methodology applies and develops general ideas on chance-constrained optimization [54] to the setting of the OPF under uncertainty. In section 4.1 we will provide a generic formulation of

our chance-constrained OPF problem that is valid under the assumption of linear power flow laws and statistical independence of wind fluctuations at different sites, while using control law (3.11) to specify standard generation response to wind fluctuations. Under the additional assumption that the fluctuations are normally distributed, in section 4.2 this formulation is reduced to a deterministic convex optimization problem, more precisely, a second-order cone program (SOCP) [21, 35]; an efficient computational implementation is discussed in section 4.3.

Many of our assumptions are not restrictive and allow natural generalizations. Using techniques from [54], one can relax the assumption of normally distributed wind sources. For example, using only the mean and variance of output at each wind farm, one can use Chebyshev's inequality to obtain a similar though more conservative formulation. And following [54] one can also obtain convex approximations to (3.8) which are tighter than Chebyshev's inequality, for a large number of empirical distributions discussed in the literature. The data-robust version of our algorithm discussed in [19] provides a methodologically sound (and computationally efficient) means to protect against data and model errors. Additionally out-of-sample experiments (section 5.5) involving the controls computed with the nominal approach (first to investigate the effect of parameter estimation errors in the normally distributed case, and, second, to gauge the impact of non-normal wind distributions) indicate robustness.

# Chapter 4

## Solving the chance-constrained OPF problem

In this chapter we first formally describe the chance-constrained optimal power flow. We then make use of the simplifying assumptions from section 3.5 to convert the chance constraints (which include probability statements) into a form usable by optimization software. The resulting convex optimization problem, though efficiently solvable in theory, raises numerical difficulties for off-the-shelf optimization algorithms on large grids. In the final section we describe a simple procedure to overcome the numerical difficulties and quickly solve the the CC-OPF on real grids with thousands of buses and lines.

### 4.1 Chance-constrained optimal power flow: formal expression

Recall that we assume each wind source  $k$  produces power of the form  $\mu_k + \omega_k$ , where  $\mu_k$  is constant and  $\omega_k$ <sup>1</sup> is a zero mean independent random variable with standard deviation  $\sigma_k$ . Eqs. (3.11) explain the affine control, where the  $\alpha_k$  are decision variables which control the

---

<sup>1</sup>Further recall that variables in bold denote random/fluctuating quantities.

proportional response of each generator to the total difference in power from the forecast. We also use the usual power output decision variables  $\bar{p}_k$ , which the standard OPF (3.2) uses as well. In our case, the bar indicates that these are the average power outputs, representing what would be generated in the case where the wind did not fluctuate at all from the forecasted average.

### 4.1.1 Preliminaries

For convenience of notation, we let every bus have a wind power source but set the wind power to zero for the other buses. Formally, where  $n$  is the number of buses, for  $k \notin \mathcal{W}$  we write  $\mu_k = \sigma_k = 0$  (so that  $\omega_k = 0$ ), thereby obtaining vectors  $\mu, \sigma, \omega \in \mathbb{R}^n$ . Likewise, we extend  $\bar{p}$  and  $\alpha$  to vectors in  $\mathbb{R}^n$  by writing  $\bar{p}_k = \alpha_k = 0$  whenever  $k \notin \mathcal{G}$ .

**Definition.** We say that the pair  $\bar{p}, \alpha$  is *viable* if the generator outputs under control law (3.11) exactly match total demand under any configuration of uncertain outputs.

The following simple result characterizes this condition as well as other basic properties of the affine control. Here and below,  $e \in \mathbb{R}^n$  is the vector of all ones.

**Lemma 4.1.1.** Under the control law (3.11) the net output of bus  $k$  equals

$$\bar{p}_k + \mu_k - d_k + \omega_k - \alpha_k(e^T \omega), \quad (4.1)$$

and thus the (stochastic) power flow equations can be written as

$$B\theta = \bar{p} + \mu - d + \omega - (e^T \omega)\alpha. \quad (4.2)$$

Consequently, the pair  $\bar{p}, \alpha$  is viable if and only if

$$\begin{aligned} \sum_{k \in \mathcal{V}} (\bar{p}_k + \mu_k - d_k) &= 0, \\ \sum_k \alpha_k &= 1. \end{aligned} \quad (4.3)$$

*Proof.* Eq. (4.1) follows by definition of the  $\bar{p}$ ,  $\mu$ ,  $d$  vectors and the control law. Thus Eq. (4.2) holds. By Lemma 3.2.1 from Eq. (4.2) one gets that  $\bar{p}, \alpha$  is viable iff

$$\begin{aligned} 0 &= \sum_{k=1}^n (\bar{p}_k - (e^T \boldsymbol{\omega}) \alpha_k + \mu_k + \boldsymbol{\omega}_k - d_k) \\ &= \sum_k (\bar{p}_k + \mu_k - d_k), \end{aligned}$$

since by construction  $\sum_k \alpha_k = 1$ . □

Equation (4.3) can be interpreted as stating the condition that expected total generation must equal total demand; however the Lemma contains a rigorous proof of this fact.

Because the admittance matrix  $B$  is singular with rank  $n - 1$ , it is conventional to eliminate the row corresponding to the slack bus and the column corresponding to the reference angle bus to obtain a system with a unique solution. Our formulation does not include a single slack bus, but rather every bus with  $\alpha_k > 0$  is used to balance power fluctuations. For convenience we will assume that bus  $n$  is the reference bus, and furthermore neither a generator nor a wind farm bus, that is to say,  $n \notin \mathcal{G} \cup \mathcal{W}$ . We denote by  $\hat{B}$  the invertible submatrix obtained from  $B$  by removing row and column  $n$ , and write

$$\check{B} = \begin{pmatrix} \hat{B}^{-1} & 0 \\ 0 & 0 \end{pmatrix}. \quad (4.4)$$

We will use  $\check{B}$  as a sort of pseudoinverse of  $B$ . By Lemma 3.2.1 we can assume without loss of generality that  $\boldsymbol{\theta}_n = 0$ . The following simple result will be used in what follows.

**Lemma 4.1.2.** Suppose the pair  $\bar{p}, \alpha$  is viable. Then under the control law (3.11) a vector of (stochastic) phase angles is

$$\boldsymbol{\theta} = \bar{\boldsymbol{\theta}} + \check{B}(\boldsymbol{\omega} - (e^T \boldsymbol{\omega}) \boldsymbol{\alpha}), \quad \text{where} \quad (4.5)$$

$$\bar{\boldsymbol{\theta}} = \check{B}(\bar{p} + \boldsymbol{\mu} - d). \quad (4.6)$$

As a consequence,

$$\mathbb{E}_\omega[\boldsymbol{\theta}] = \bar{\boldsymbol{\theta}}, \quad (4.7)$$

and given any line  $\{k, m\}$ ,

$$\mathbb{E}_\omega[\mathbf{f}_{km}] = \beta_{km}(\bar{\theta}_k - \bar{\theta}_m). \quad (4.8)$$

Furthermore, each quantity  $\theta_k$  or  $\mathbf{f}_{km}$  is an affine function of the random variables  $\omega_k$ .

*Proof.* For convenience we rewrite system (4.2):  $B\boldsymbol{\theta} = \bar{p} + \mu - d + \boldsymbol{\omega} - (e^T \boldsymbol{\omega})\alpha$ . Since  $\bar{p}, \alpha$  is viable, this system is always feasible, and since the sum of rows of  $B$  is zero, its last row is redundant. Therefore Eq. (4.5) follows since  $\theta_n = 0$ , and Eq. (4.7) holds since  $\boldsymbol{\omega}$  has zero mean. Since  $\mathbf{f}_{km} = \beta_{km}(\theta_k - \theta_m)$  for all  $\{k, m\}$ , Eq. (4.8) holds. From this fact and Eq. (4.5) it follows that  $\boldsymbol{\theta}$  and  $\mathbf{f}$  are affine functions of  $\boldsymbol{\omega}$ .  $\square$

### 4.1.2 Formal expression

Using this result we can now give an initial formulation to our chance-constrained problem. Because this formulation includes probability statements, it is not directly usable by optimization software. In the next section we will use the assumptions from section 3.5 to transform this to a

conventional convex optimization formulation.

$$\mathbf{CC-OPF:} \quad \min E_{\omega} [c(\bar{p} - (e^T \omega)\alpha)] \quad (4.9)$$

$$\text{s.t.} \quad \sum_{g \in \mathcal{G}} \alpha_g = 1, \quad \alpha \geq 0, \quad \bar{p} \geq 0, \quad (4.10)$$

$$\sum_{k \in \mathcal{V}} (\bar{p}_k + \mu_k - d_k) = 0, \quad (4.11)$$

$$B\bar{\theta} = \bar{p} + \mu - d; \quad (4.12)$$

for all lines  $\{k, m\}$ :

$$P \left( \beta_{km}(\bar{\theta}_k - \bar{\theta}_m + [\check{\mathcal{B}}(\omega - (e^T \omega)\alpha)]_k - [\check{\mathcal{B}}(\omega - (e^T \omega)\alpha)]_m) > f_{km}^{max} \right) < \epsilon_{km}, \quad (4.13)$$

$$P \left( \beta_{km}(\bar{\theta}_k - \bar{\theta}_m + [\check{\mathcal{B}}(\omega - (e^T \omega)\alpha)]_k - [\check{\mathcal{B}}(\omega - (e^T \omega)\alpha)]_m) < -f_{km}^{max} \right) < \epsilon_{km}; \quad (4.14)$$

for all generators  $g$ :

$$P (\bar{p}_g - (e^T \omega)\alpha_g > p_g^{max}) < \epsilon_g \text{ and } P (\bar{p}_g - (e^T \omega)\alpha_g < p_g^{min}) < \epsilon_g. \quad (4.15)$$

The variables in this formulation are  $\bar{p}$ ,  $\alpha$ , and  $\bar{\theta}$ . Constraint (4.10) simply states basic conditions needed by the affine control. Constraint (4.11) is (4.3). Constraints (4.12), (4.13) and (4.14) express our chance constraint, in view of Lemma 4.1.2.

### 4.1.3 Objective

The objective function is the expected cost incurred by the stochastic generation vector

$$\mathbf{p} = \bar{p} - (e^T \omega)\alpha$$

over the varying wind power output  $\omega$ . In standard power engineering practice generation cost is convex, quadratic and separable, i.e. for any vector  $p$ ,  $c(p) = \sum_k c_k(p_k)$  where each  $c_k$  is convex quadratic. Note that for any  $g \in \mathcal{G}$  we have

$$\mathbf{p}_g^2 = \bar{p}_g^2 + (e^T \omega)^2 \alpha_g^2 - 2e^T \omega \bar{p}_g \alpha_g,$$

from which we obtain, since the  $\omega_g$  have zero mean,

$$\mathbb{E}_{\omega}[p_g^2] = \bar{p}_g^2 + \text{var}(\Omega)\alpha_g^2,$$

where “var” denotes variance and  $\Omega \doteq \sum_k \omega_k$ . It follows that the objective function can be written as

$$\mathbb{E}_{\omega}[c(\mathbf{p})] = \sum_{g \in \mathcal{G}} \{c_{g2} (\bar{p}_g^2 + \text{var}(\Omega)\alpha_g^2) + c_{g1}\bar{p}_g + c_{g0}\} \quad (4.16)$$

where  $c_{g2} \geq 0$  for all  $g \in \mathcal{G}$ . Consequently the objective function is convex quadratic, as a function of  $\bar{p}$  and  $\alpha$ .

#### 4.1.4 Variance of the line flow $f_{km}$

The above formulation is the formal statement for our optimization problem. Even though its objective is convex in cases of interest, the formulation is not in a form that can be readily exploited by standard optimization algorithms. Below we will provide an efficient approach to solve relevant classes of problems with the above form; prior to that we need a technical result. We will employ the following notation:

- For  $1 \leq k, m \leq n$  let  $\pi_{km}$  denote the  $k, m$  entry of the matrix  $\check{B}$  given above, that is,

$$\pi_{km} = \begin{cases} [\hat{B}^{-1}]_{km}, & k < n, \\ 0 & \text{otherwise.} \end{cases} \quad (4.17)$$

- Given  $\alpha$ , for  $1 \leq k \leq n$  write

$$\delta_k \doteq [\check{B}\alpha]_k = \begin{cases} \sum_{m=1}^{n-1} \pi_{km}\alpha_m, & k < n, \\ 0 & \text{otherwise.} \end{cases} \quad (4.18)$$

We use  $\pi$  as simpler notation for  $\check{B}$ , and  $\delta$  is the “solution” to the system  $B\delta = \alpha$  using the

pseudoinverse  $\check{\mathbf{B}}$ .

**Lemma 4.1.3.** Assume that the  $\omega_k$  are independent random variables. Given  $\alpha$ , for any line  $\{k, m\}$ ,

$$\text{var}(\mathbf{f}_{km}) = \beta_{km}^2 \sum_{l \in \mathcal{W}} \sigma_l^2 (\pi_{kl} - \pi_{ml} - \delta_k + \delta_m)^2. \quad (4.19)$$

*Proof.* Using  $\mathbf{f}_{km} = \beta_{km}[\boldsymbol{\theta}_k - \boldsymbol{\theta}_m]$  and eq. (4.5) we have that

$$\begin{aligned} \mathbf{f}_{km} - \mathbb{E}_{\boldsymbol{\omega}}[\mathbf{f}_{km}] &= \beta_{km}([\check{\mathbf{B}}(\boldsymbol{\omega} - (e^T \boldsymbol{\omega})\alpha)]_k - [\check{\mathbf{B}}(\boldsymbol{\omega} - (e^T \boldsymbol{\omega})\alpha)]_m) \\ &= \beta_{km} \left( [\check{\mathbf{B}}\boldsymbol{\omega}]_k - [\check{\mathbf{B}}\boldsymbol{\omega}]_m - (e^T \boldsymbol{\omega})\delta_k + (e^T \boldsymbol{\omega})\delta_m \right) \\ &= \beta_{km} \sum_{l \in \mathcal{W}} (\pi_{kl} - \pi_{ml} - \delta_k + \delta_m) \omega_l, \end{aligned}$$

since by convention  $\omega_k = 0$  for any  $k \notin \mathcal{W}$ . The result now follows.  $\square$

*Remark.* Lemma 4.1.3 holds for any distribution of the  $\omega_k$  so long as independence is assumed. Similar results are easily obtained for higher-order moments of the  $\mathbf{f}_{km}$ .

## 4.2 Formulating the chance-constrained problem as a conic program

In deriving the above formulation (4.9)-(4.15) for the CC-OPF we assumed that the  $\omega_k$  random variables have zero mean. To obtain an efficient solution procedure we will additionally assume that they are (a) independent and (b) normally distributed, as described in section 3.5. Under these additional assumptions, since the  $\mathbf{f}_{km}$  are affine functions of the  $\omega_k$ , there is a simple restatement of the chance-constraints (4.13), (4.14), and (4.15) in a computationally practicable form. See [54] for a general treatment of linear inequalities with stochastic coefficients. For any real  $0 < r < 1$  we write  $\eta(r) = \Phi^{-1}(1 - r)$ , where  $\Phi$  is the cumulative distribution function of the standard normal distribution.

**Lemma 4.2.1.** Let  $\bar{p}$ ,  $\alpha$  be viable. Assume that the  $\omega_k$  are normally distributed and independent. Then the following hold:

- For any line  $\{k, m\}$ ,  $P(\mathbf{f}_{km} > f_{km}^{max}) < \epsilon_{km}$  and  $P(-\mathbf{f}_{km} > f_{km}^{max}) < \epsilon_{km}$  iff

$$\beta_{km}|\bar{\theta}_k - \bar{\theta}_m| \leq f_{km}^{max} - \eta(\epsilon_{km}) \left[ \beta_{km}^2 \sum_{l \in \mathcal{W}} \sigma_l^2 (\pi_{kl} - \pi_{ml} - \delta_k + \delta_m)^2 \right]^{1/2} \quad (4.20)$$

where as before  $\bar{\theta} = \check{\mathcal{B}}(\bar{p} + \mu - d)$  and  $\delta = \check{\mathcal{B}}\alpha$ .

- For any generator  $g$ ,  $P(\bar{p}_g - (e^T \omega)\alpha_g > p_g^{max}) < \epsilon_g$  and  $P(\bar{p}_g - (e^T \omega)\alpha_g < p_g^{min}) < \epsilon_g$  iff

$$p_g^{min} + \eta(\epsilon_g) \left( \sum_{k \in \mathcal{W}} \sigma_k^2 \right)^{1/2} \leq \bar{p}_g \leq p_g^{max} - \eta(\epsilon_g) \left( \sum_{k \in \mathcal{W}} \sigma_k^2 \right)^{1/2}. \quad (4.21)$$

*Proof.* By Lemma 4.1.2,  $\mathbf{f}_{km}$  is an affine function of the  $\omega_k$ ; under the assumption it follows that  $\mathbf{f}_{km}$  is itself normally distributed. Thus,  $P(\mathbf{f}_{km} > f_{km}^{max}) < \epsilon_{km}$  iff

$$\mathbb{E}_\omega[\mathbf{f}_{km}] + \eta(\epsilon_{km}) \text{var}(\mathbf{f}_{km}) \leq f_{km}^{max}, \quad (4.22)$$

and similarly,  $P(\mathbf{f}_{km} < -f_{km}^{max}) < \epsilon_{km}$  iff

$$\mathbb{E}_\omega[\mathbf{f}_{km}] - \eta(\epsilon_{km}) \text{var}(\mathbf{f}_{km}) \geq -f_{km}^{max}. \quad (4.23)$$

Lemma 4.1.2 gives  $\mathbb{E}_\omega[\mathbf{f}_{km}] = \beta_{km}(\bar{\theta}_k - \bar{\theta}_m)$  while by Lemma 4.1.3,  $\text{var}(\mathbf{f}_{km}) = \beta_{km}^2 \sum_{l \in \mathcal{W}} \sigma_l^2 (\pi_{kl} - \pi_{ml} - \delta_k + \delta_m)^2$ . Substituting these values into (4.22) and (4.23) yields (4.20). The proof of (4.21) is similar.  $\square$

*Remark.* Eq. (4.20) highlights the difference between e.g. our chance constraint for lines, which requires that  $P(\mathbf{f}_{km} > f_{km}^{max}) < \epsilon_{km}$  and that  $P(\mathbf{f}_{km} < -f_{km}^{max}) < \epsilon_{km}$ , and the stricter requirement that  $P(|\mathbf{f}_{km}| > f_{km}^{max}) < \epsilon_{km}$  which amounts to

$$P(\mathbf{f}_{km} > f_{km}^{max}) + P(\mathbf{f}_{km} < -f_{km}^{max}) < \epsilon_{km}. \quad (4.24)$$

Unlike our requirement, which is captured by (4.20), the stricter condition (4.24) does not admit a compact statement.

We can now present a formulation of our chance-constrained optimization as a convex optimization problem, on variables  $\bar{p}, \alpha, \bar{\theta}, \delta$ , and  $s$ . We will assume in what follows that for all lines  $\{k, m\}$ ,  $\epsilon_{km} < 1/2$ , so that  $\eta(\epsilon_{km}) > 0$ .

$$\min \sum_{k \in \mathcal{G}} \left\{ c_{k2} \left( \bar{p}_k^2 + \alpha_k^2 \sum_m \sigma_m^2 \right) + c_{k1} \bar{p}_k + c_{k0} \right\}; \quad (4.25)$$

$$\text{for } 1 \leq k \leq n-1: \quad \sum_{m=1}^{n-1} \hat{B}_{km} \delta_m = \alpha_k; \quad (4.26)$$

$$\text{for } 1 \leq k \leq n-1: \quad \sum_{m=1}^{n-1} \hat{B}_{km} \bar{\theta}_m - \bar{p}_k = \mu_k - d_i; \quad (4.27)$$

$$\sum_k \alpha_k = 1, \quad \alpha \geq 0, \quad \bar{p} \geq 0; \quad (4.28)$$

$$\bar{p}_n = \alpha_n = \delta_n = \bar{\theta}_n = 0; \quad (4.29)$$

$$\beta_{km} |\bar{\theta}_k - \bar{\theta}_m| + \beta_{km} \eta(\epsilon_{km}) s_{km} \leq f_{km}^{max} \quad \forall \{k, m\}; \quad (4.30)$$

$$\left[ \sum_{l \in \mathcal{W}} \sigma_l^2 (\pi_{kl} - \pi_{ml} - \delta_k + \delta_m)^2 \right]^{1/2} - s_{km} \leq 0 \quad \forall \{k, m\}; \quad (4.31)$$

$$-\bar{p}_g + \eta(\epsilon_g) \left( \sum_{k \in \mathcal{W}} \sigma_k^2 \right)^{1/2} \leq -p_g^{min} \quad \forall g \in \mathcal{G}; \quad (4.32)$$

$$\bar{p}_g + \eta(\epsilon_g) \left( \sum_{k \in \mathcal{W}} \sigma_k^2 \right)^{1/2} \leq p_g^{max} \quad \forall g \in \mathcal{G}; \quad (4.33)$$

In this formulation, the variables  $s_{km}$  are auxiliary and introduced to facilitate the discussion below – since  $\eta(\epsilon_{km}) \geq 0$  without loss of generality (4.31) will hold as an equality. Constraints (4.31), (4.32), and (4.33) are *second-order cone* inequalities [21]. A problem of the above form is solvable in polynomial time using well-known methods of convex optimization; several commercial software tools such as CPLEX [6], Gurobi [10], MOSEK [7] and others are available. Constraint (4.26) is equivalent to  $\delta_k = \sum_{m=1}^{n-1} \pi_{km} \alpha_m$  (as we did in (4.18)), however the  $\pi_{km}$  can be seen

to be all nonzero, whereas  $\hat{B}$  is very sparse for typical grids. Constraint (4.31) can be relatively dense – the sum has a term for each wind farm. As a percentage of the total number of buses, however, this can be expected to be small.

### 4.3 Solving the conic program

Even though optimization theory guarantees that the above problem is efficiently solvable, experimental testing shows that in the case of large grids (thousands of lines) the problem proves challenging. For example, in the Polish 2003-2004 winter peak case,<sup>2</sup> we have 2746 buses, 3514 lines and 8 wind farms, and CPLEX [6] reports (after pre-solving) 36625 variables and 38507 constraints, of which 6242 are conic. On this problem, a recent version of CPLEX on a modern workstation<sup>3</sup> ran for 3392 seconds and was unable to produce a feasible solution. On the same problem Gurobi reported “numerical trouble” after 31.1 cpu seconds and stopped.

In fact, all of the commercial solvers [6, 7, 10] we experimented with reported numerical difficulties with problems of this size. Anecdotal evidence indicates that the primary cause for these difficulties is not simply the size, but also to a large degree *numerics* in particular poor conditioning due to the entries in the matrix  $B$ . These are susceptances, which are inverses of reactances, and often take values in a wide range.

To address this issue we implemented an effective algorithm for solving problem (4.25)-(4.33). For brevity we will focus on constraints (4.31) ((4.32) and (4.33) are similarly handled). For a line  $\{k, m\}$  define

$$C_{km}(\delta) \doteq \left( \sum_{l \in \mathcal{W}} \sigma_l^2 (\pi_{kl} - \pi_{ml} - \delta_k + \delta_m)^2 \right)^{1/2}.$$

Constraint (4.31) can thus be written as  $C_{km}(\delta) \leq s_{km}$ . For completeness, we state the following result:

---

<sup>2</sup>Available with MATPOWER [72]

<sup>3</sup>A machine with two Intel Xeon CPU X5570 2.93GHz chips (total of 8 cores) and 64GB RAM. CPLEX ran on 16 parallel threads, making use of “hyperthreading.”

**Lemma 4.3.1.** Constraint (4.31) is equivalent to the infinite set of linear inequalities

$$C_{km}(\hat{\delta}) + \frac{\partial C_{km}(\hat{\delta})}{\partial \delta_k}(\delta_k - \hat{\delta}_k) + \frac{\partial C_{km}(\hat{\delta})}{\partial \delta_m}(\delta_m - \hat{\delta}_m) \leq s_{km}, \quad \forall \hat{\delta} \in \mathbb{R}^n \quad (4.34)$$

Constraints (4.34) express the *outer envelope* of the set described by (4.31) [21]. Any vector  $\delta \in \mathbb{R}^n$  which satisfies (4.31) (for a given choice of  $s_{km}$ ) is guaranteed to satisfy (4.34). Thus a finite subset of the inequalities (4.34), used instead of (4.31), will give rise to a relaxation of the optimization problem and thus a lower bound on the optimal objective value. Given Lemma 4.3.1 there are two ways to proceed, both motivated by the observation that at  $\delta^* \in \mathbb{R}^n$ , the most constraining inequality from among the set (4.34) (that is to say, the one whose left-hand side evaluated at  $\delta^*$  is largest) is that obtained by choosing  $\hat{\delta} = \delta^*$ .

First, one can use inequalities (4.34) as *cutting-planes* in the context of the ellipsoid method [36], obtaining a polynomial-time algorithm. A different way to proceed yields a numerically practicable algorithm, described below. For a classical reference, see [40]. Denote by  $F(\bar{p}, \alpha)$  the objective function in Eq. (4.25).

**Algorithm 4.1:** Cutting plane procedure

The linear “master” system  $A(\bar{p}, \alpha, \delta, \theta, s)^T \geq b$  is defined to include constraints (4.26)-(4.30)

**Iterate**

- 1 Set  $(\bar{p}^*, \alpha^*, \delta^*, \theta^*, s^*) \leftarrow \operatorname{argmin}\{F(\bar{p}, \alpha) : A(\bar{p}, \alpha, \delta, \theta, s)^T \geq b\}$
- 2 **if** all conic constraints are satisfied by  $(\bar{p}^*, \alpha^*, \delta^*, \theta^*, s^*)$  **then** Stop
- 3 **else if** all chance constraints are satisfied by  $(\bar{p}^*, \alpha^*)$  **then** Stop
- 4 Add to the master system the outer inequality (4.34) arising from the constraint (4.31) which is most violated

**End**

As the algorithm iterates the master system represents a valid relaxation of the conic program (4.26)-(4.31); thus the objective value of the solution computed in Step 1 is a valid lower bound

on the value of problem. Each problem solved in Step 1 is a linearly constrained, convex quadratic program. Computational experiments involving large-scale realistic cases show that the algorithm is robust and rapidly converges to an optimum.

Note that Step 3 is not redundant. Recall that our formulation includes the variables  $s_{km}$ ; each such variable is lower bounded (via equation (4.31)) by the variance of flow on line  $\{k, m\}$ . However in the cutting-plane algorithm we do not make explicit use of equation (4.31) – indeed, we progressively approximate it by means of cutting planes. At a typical intermediate iteration the current solution will still fail to satisfy (4.31), i.e. we may have, for some line (or lines)  $\{k, m\}$ ,

$$\left[ \sum_{l \in \mathcal{W}} \sigma_l^2 (\pi_{kl} - \pi_{ml} - \delta_k^* + \delta_m^*)^2 \right]^{1/2} > s_{km}^*, \quad (4.35)$$

(where “\*” indicates the current values of the variables) which would normally indicate that the algorithm has not terminated because condition (4.31) has not been adequately represented by the cutting-planes. However, in later stages of the algorithm it may be the case that (4.35) holds for some lines, and yet the pair  $(\bar{p}^*, \alpha^*)$  already satisfies the chance constraints for all lines. This situation will arise if in the current solution some of the  $s_{km}^*$  quantities are artificially low, that is to say, they could be increased so as to match the left-hand side of (4.35) without jeopardizing feasibility of the rest of the solution, that is to say, constraint (4.30). The underlying dynamic is that the conic constraints (4.31) are still somewhat loosely (outer-)approximated by a set of linear inequalities, resulting in a range of possible values for a particular variable  $s_{km}$ . One such value will be picked by the underlying optimization solver; and as just argued this value may be artificially low. A direct way to check that  $(\bar{p}^*, \alpha^*)$  satisfies the chance constraint for a given line  $\{k, m\}$ , is straightforward – the flows  $f_{km}$  are normally distributed (as noted in Lemma 4.2.1) and their means and variances can be directly computed from  $(\bar{p}^*, \alpha^*)$ , and so we can directly compute the probability that, under  $(\bar{p}^*, \alpha^*)$ , any given  $f_{km}$  exceeds its corresponding line limit (additionally, this procedure computes said probability, a useful output parameter). If  $(\bar{p}^*, \alpha^*)$  indeed satisfies all chance constraints, it is optimal, since it attains a cost that matches a lower

bound to the overall problem (i.e. it attains the value of the current cutting-plane relaxation).

### 4.3.1 Example of typical performance of the cutting plane algorithm

In our implementation termination is declared in Step 2 or Step 3 when the corresponding constraint violation is less than  $10^{-6}$ . Table 4.1 displays typical performance of the cutting-plane algorithm on (comparatively more difficult) large problem instances. In the table, 'Polish1'- 'Polish3' are the three Polish cases included in MATPOWER [72] (in Polish1 we increased loads by 30%). All Polish cases have uniform random costs on  $[0.5, 2.0]$  for each generator and ten arbitrarily chosen wind sources. The average wind power penetration for the four cases is 8.8%, 3.0%, 1.9%, and 1.5%. 'Iterations' is the number of linearly-constrained subproblems solved before the algorithm converges. 'Barrier iterations' is the total number of iterations of the barrier algorithm in CPLEX over all subproblems, and 'Time' is the total (wallclock) time required by the algorithm. Line tolerances are set to two standard deviations and generator tolerances three standard deviations (tail probabilities 0.0228 and 0.00135, respectively). These instances all prove unsolvable if directly tackled by CPLEX or Gurobi. A more challenging test involving the Polish grid is reported in Section 5.6.

Table 4.1: Performance of cutting-plane method on a number of large cases.

Case	Buses	Generators	Lines	Time (s)	Iterations	Barrier iterations
BPA	2209	176	2866	5.51	2	256
Polish1	2383	327	2896	13.64	13	535
Polish2	2746	388	3514	30.16	25	1431
Polish3	3120	349	3693	25.45	23	508

Table 4.2 provides additional, typical numerical performance for the cutting-plane algorithm on an instance of the Polish grid model. Each row of Table 4.2 shows, for selected iterations, the current objective value as well as the maximum percentage-wise feasibility error. This quantity is defined as follows: we maximize, over all lines  $\{k, m\}$ , the ratio of the amount by which the left-hand side of (4.30) exceeds the right-hand side, to the right-hand side. The total run-time

was 25 seconds. Note the “flatness” of the objective. This makes the problem nontrivial – the challenge is to find a *feasible* solution (with respect to the chance constraints); at the onset of the algorithm the computed solution is quite infeasible and it is this attribute that is quickly improved by the cutting-plane algorithm.

Table 4.2: Typical convergence behavior of cutting-plane algorithm on a large instance.

Iteration	Max rel. error	Objective
1	1.2e-1	7.0933e6
4	1.3e-3	7.0934e6
7	1.9e-3	7.0934e6
10	1.0e-4	7.0964e6
12	8.9e-7	7.0965e6

We note the (typical) small number of iterations needed to attain numerical convergence. Thus at termination only a very small number of conic constraints (4.31) have been incorporated into the master system. This validates the expectation that only a small fraction of the conic constraints in the CC-OPF are active at optimality. The cutting-plane algorithm can be viewed as a procedure that opportunistically discovers these constraints.

# Chapter 5

## Numerical experiments

Here we will describe qualitative aspects of our affine control on small systems; in particular we focus on the contrast between the standard OPF and the CC-OPF on problematic features that can arise because of fluctuating wind sources and on out-of-sample testing of the CC-OPF solution, including the analysis of non-normal distributions. Some of our tests involve the BPA grid and the Polish grid, which are large; we present additional sets of tests to address the scalability of our solution methodology to the large cases.

Above (see Eq. (3.2)) we introduced the so-called standard OPF method for setting traditional generator output levels. When renewables are present, the natural extension of this approach would make use of some fixed estimate of output (e.g. mean output) and to handle fluctuations in renewable output through the same method used to deal with changes in load: ramping output of traditional generators up or down in proportion to the net increase or decrease in renewable output. This feature could seamlessly be handled using today's control structure, with each generator's output adjusted at a fixed (preset) rate. For the sake of simplicity, in the experiments below we assume that all ramping rates are equal. That is,  $\alpha_g = 1/n_g \forall g \in \mathcal{G}$  where  $n_g$  is the number of generators.

Different assumptions on these fixed rates will likely produce different numerical results; however, this general approach entails an inherent weakness. The key point here is that mean

generator output levels *as well as* in particular the ramping rates would be chosen *without* considering the stochastic nature of the renewable output levels. Our experiments are designed to highlight the limitations of this “risk-unaware” approach. In contrast, our CC-OPF produces control parameters (the  $\bar{p}$  and the  $\alpha$ ) that are risk-aware and, implicitly, also topology-aware – in the sense of network proximity to wind farms.

As a technical point, we note that generator adjustment under the standard OPF can be viewed as a special case of our affine control mechanism, but with fixed  $\alpha$  values. Above, (Lemmas 4.1.2 and 4.1.3) we have provided expressions for the mean and standard deviation of the flow on any given line, under the independence assumption for the  $\omega_k$ . It follows that under the normally-distributed assumption, for any given vectors  $\bar{p}$  and  $\alpha$  we can compute the probability that any given line  $\{k, m\}$  is overloaded (both under the CC-OPF and the standard OPF).

## 5.1 Failure of standard OPF

We first consider the IEEE 118-bus model with a quadratic cost function, and four sources of wind power added at arbitrary buses to meet 5% of demand in the case of average wind. The standard OPF solution is safely within the thermal capacity limits for all lines in the system. Then we account for fluctuations in wind assuming normally-distributed and site-independent fluctuations with standard deviations set to 30% of the respective means. The results, which are shown in Fig. 5.1, illustrate that under the standard OPF five lines (marked in red) frequently become overloaded, exceeding their limits 8% or more of the time. This situation translates into an unacceptably high risk of failure for any of the five red lines. This problem occurs for grids of all sizes; similar results hold on the 2746-bus Polish grid (from MATPOWER [72]). In this case, after scaling up all loads by 10% to simulate a more highly stressed system, we added wind power to ten buses for a total of 2% penetration. The standard solution results in six lines exceeding their limits over 45% of the time, and in one line over 10% of the time. For an additional and

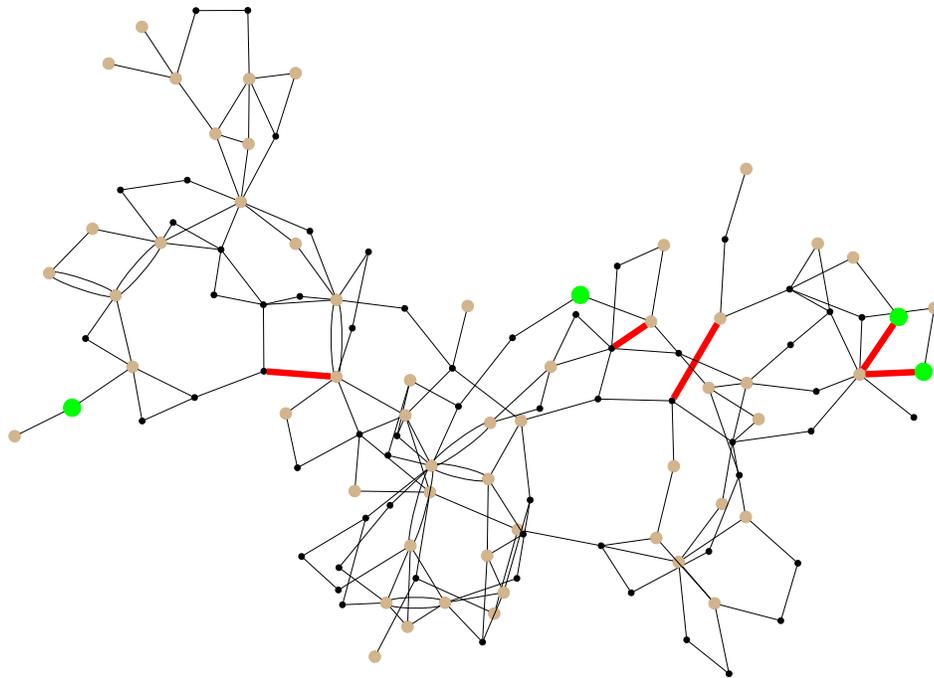


Figure 5.1: 118-bus case with four wind farms (green dots; brown are generators, black are loads). Shown is the standard OPF solution against the average wind case with penetration of 5%. Standard deviations of the wind are set to 30% of the respective average cases. Lines in red exceed their limit 8% or more of the time.

similar experiment using the Polish grid see Section 5.6.

## 5.2 Cost of reliability under high wind penetration

If we maintain the current operational paradigm, congestion of transmission lines may force temporary shutdown of wind farms even during times of high wind. Our methodology suggests, as an alternative solution to curtailment of wind power, an appropriate reconfiguration of standard generators. If successful, this solution can use the available wind power without curtailment, and

thus result in cheaper operating costs.

As a (crude) proxy for curtailment, we perform the following experiment, which considers different levels of renewable penetration. Here, the mean power outputs of the wind sources are kept in a fixed proportion to one another and proportionally scaled so as to vary total amount of penetration, and likewise with the standard deviations. Under a high penetration level (e.g. 30%), the standard OPF solution may yield significant probability of line overloads. In order to obtain comparisons with CC-OPF, we perform experiments that involve two controls: (1) increasing line limits by some proportion and (2) decreasing wind penetration. Assuming zero cost for wind power, the difference in cost for a high-penetration CC-OPF solution and the low-penetration standard solution represents the savings produced by our model (generously, given the line limit increase).

We performed experiments involving the 39-bus case. For this test the standard cost function associated with this case (provided with MATPOWER), which is  $\sum_{g \in G} (0.01 p_g^2 + 0.3 p_g)$  (we are ignoring constant terms). The following table summarizes the outcome of this experiment.

Table 5.1: Impact of penetration, 39-bus case

OPF type	Penetration	Line limit boost	Overload prob.	Cost
STD	30 %	10 %	21.8 %	$2.05 \times 10^4$
STD	5 %	0 %	0.0 %	$3.72 \times 10^4$
CC	30 %	0 %	0.3 %	$2.16 \times 10^4$
CC	40 %	0 %	0.5 %	$1.80 \times 10^4$

In this table, “OPF type” refers to either the standard (“STD”) or the CC-OPF, “Line limit boost” is the percentage by which line limits are increased, and “Overload prob.” is the maximum line overload probability. Thus, in the case of 30% penetration, the standard OPF experiences high probability of line overloads, even with a 10% increase in line limits (two lines are overloaded with very high probability). Reducing penetration to 5% removes this problem, but at a major cost increase. In contrast, even under 40% penetration, the CC-OPF achieves a low overload probability and a significant cost decrease.

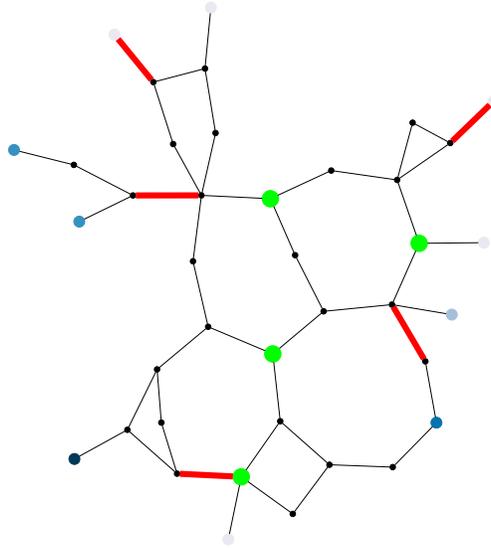


Figure 5.2: 39-bus case. Red lines indicate high probability of flow exceeding the limit under the standard OPF solution. Generators are shades of blue, with darker shades indicating greater absolute difference between the chance-constrained solution and the standard solution.

### 5.3 Non-locality

We have established that under fluctuating power generation, some lines may exceed their flow limits an unacceptable fraction of the time. Is there a simple solution to this problem, for instance, by carefully adjusting (a posteriori of the standard OPF) the outputs of the generators near the violated lines? The answer is no. Power systems exhibit significant non-local behavior. Consider Fig. 5.2. In this example, the major differences in generator outputs between the standard OPF solution and our CC-OPF model's solution are not obviously associated with the different line violations. Notice, that non-locality of the power flow response is not specific to the CC-OPF, and it may also be seen in experiments with the standard OPF testing sensitivity of power flows, for example to generation shift factors. However, our point here is that non-locality caused by stochastic factors, accounted for in the CC-OPF, may be of a new type which does not allow intuitive explanation in terms of the standard OPF. In general, it seems that it would be difficult to by-pass the CC-OPF and make small local adjustments to relieve the stressed lines. On the positive side, even though the CC-OPF is not local and requires a centralized computation, it is also only slightly more difficult than the standard OPF in terms of implementation.

## 5.4 Increasing penetration

Current planning for the power system in the United States calls for 30% of wind energy penetration by 2030 [4]. Investments necessary to achieve this ambitious target may focus on both software (improving operations) and hardware (building new lines, sub-stations, etc), with the former obviously representing a much less expensive and thus economically attractive option. Our CC-OPF solution contributes toward this option. A natural question that arises concerns the maximum level of penetration one can safely achieve by upgrading from the standard OPF to our CC-OPF.

To answer the question we consider the 39-bus New England system (from [72]) case with four wind generators added, and line flow limits scaled by .7 to simulate a heavily loaded system. The quadratic cost terms (the  $c_{i2}$  coefficients in (4.16), which are the leading coefficients for the quadratic functions  $c_i$ ) are set to  $\text{rand}(0,1) + .5$ . We fix the four wind generator average outputs in a ratio of 5/6/7/8 and standard deviations at 30% of the mean. We first solve our model using  $\epsilon = .02$  for each line and assuming zero wind power, and then increase total wind output until the optimization problem becomes infeasible. See Fig. 5.3. This experiment illustrates that at least for the model considered, the 30% of wind penetration with rather strict probabilistic guarantees enforced by our CC-OPF may be feasible, but in fact lies rather close to the maximum possible. To push penetration beyond the threshold is impossible without upgrading lines and investing in other (not related to wind farms themselves) hardware.

## 5.5 Out-of-sample tests

We now study the performance of the control computed using the CC-OPF when there are errors in the underlying distribution of wind power. We consider two types of errors: (1) the true distribution is non-normal but our normal fit is “close” in an appropriate sense, and (2) the true distribution is normal but with different mean or standard deviation. The experiments in this section use as data the BPA grid, which as noted before has 2209 buses and 2866 lines, and

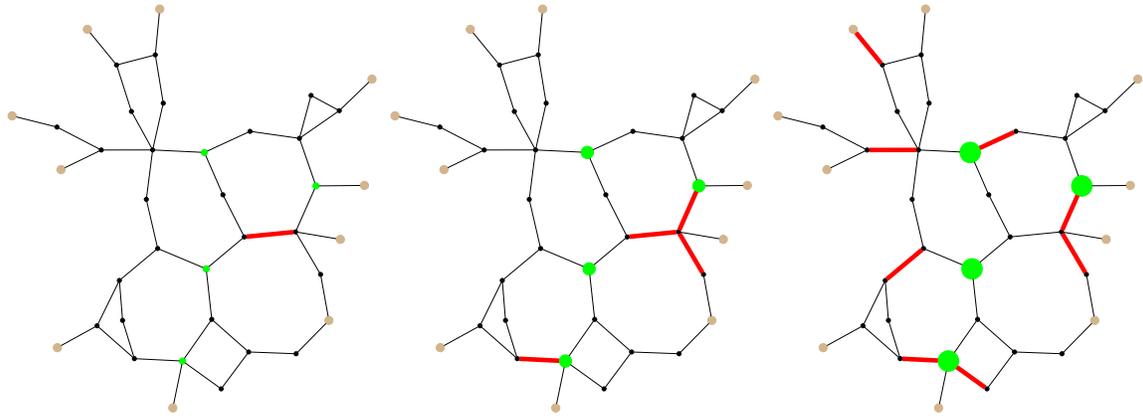


Figure 5.3: 39-bus case with four wind farms (green dots; brown are generators, black are loads). Lines in red are at the maximum of  $\epsilon = .02$  chance of exceeding their limit. The three cases shown are left to right: .1%, 8%, and 30% average wind penetration. With penetration beyond 30% the problem becomes infeasible.

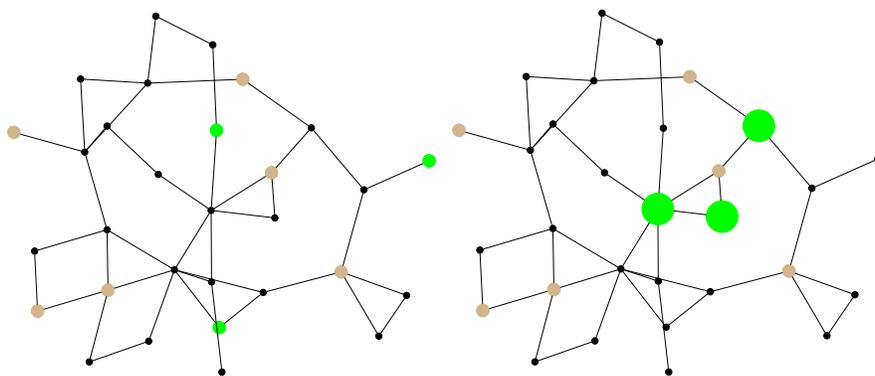


Figure 5.4: 30-bus case with three wind farms. The case on the left supports only up to 10% before becoming infeasible, while the one on the right is feasible up to 55% penetration.

collected wind data, altogether constituting a realistic test-case.

We first consider the non-normal case, using the following probability distributions, all with fatter tails than the normal distribution: (1) Laplace, (2) logistic, (3) Weibull (three different shapes), (4) t location-scale with 2.5 degrees of freedom, and (5) Cauchy. For the Laplace and logistic distributions, we simply match the mean and standard deviation. For the Weibull distribution, we consider shape parameters  $k = 1.2, 2, 4$  and choose the scale parameter to match the standard deviations. We then translate to match means. For the t distribution, we fix 2.5 degrees of freedom and then choose the location and scale to match mean and standard deviation. For the Cauchy distribution, we set the location parameter to the mean and choose the scale parameter so as to match the 95th percentiles.

We use our model and solve under the normal distribution assumption, seeking a solution which results in no line violations for cases within two standard deviations of the mean, i.e. a maximum of about 2.27% chance of exceeding the limit. We then perform Monte Carlo tests drawing from the above distributions to determine the actual rates of violation. See Fig. 5.5. The worst-performer is the highly-asymmetric (and perhaps unreasonable) Weibull with shape parameter 1.2, which approximately doubles the desired maximum chance of overload. Somewhat surprisingly, the fat-tailed logistic and Student's t distributions result in a maximum chance of overload significantly less than desired, suggesting that our model is too conservative in these cases.

Next we consider the normally distributed case with errors. We solve with nominal values for the mean and standard deviation of wind power. We then consider the rate of violation after scaling all means and standard deviations (separately). While the solution is sensitive to errors in the mean forecast, the sensitivity is well-behaved. With a desired safety level of  $\epsilon = 2.27\%$  for each line, an error in the mean of 25% results in a maximum 15% chance of exceeding the limit. The solution is quite robust to errors in the standard deviation forecast, with a 25% error resulting in less than 6% chance of overload. See Fig. 5.6.

Distribution	Max. prob. violation
Normal	0.0227
Laplace	0.0297
logistic	0.0132
Weibull, $k = 1.2$	0.0457
Weibull, $k = 2$	0.0355
Weibull, $k = 4$	0.0216
t location-scale, $\nu = 2.5$	0.0165
Cauchy	0.0276

Figure 5.5: Maximum probability of overload for out-of-sample tests. These are a result of Monte Carlo testing with 10,000 samples on the BPA case, solved under assuming normally distributed fluctuations and desired maximum chance of overload at 2.27%.

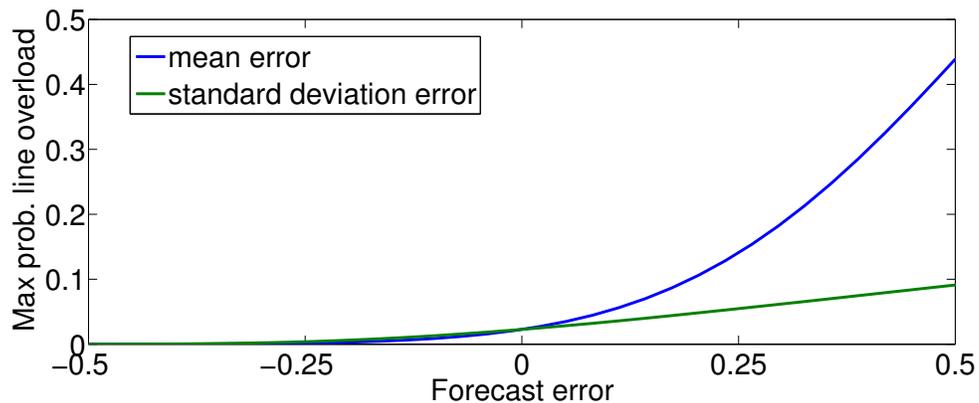


Figure 5.6: BPA case solved with average penetration at 8% and standard deviations set to 30% of mean. The maximum probability of line overload desired is 2.27%, which is achieved with 0 forecast error on the graph. Actual wind power means are then scaled according to the x-axis and maximum probability of line overload is recalculated (blue). The same is then done for standard deviations (green).

## 5.6 Scalability

As an additional experiment illustrating scalability of the approach, we studied the Polish national grid (obtained from MATPOWER as explained above) under simulated 20% renewable penetration spread over 50 wind farms, co-located with the 50 largest generators. This co-location should lessen the risk associated with renewable fluctuation (which should be partially “absorbed” by the co-located generators). Figure 5.7 studies the resulting risk exposure under the standard OPF and the CC-OPF. The chart shows the number of lines that attain several levels of overload probability. The performance of the standard OPF is unacceptable: it would lead to frequent tripping of at least four lines. In contrast, under the CC-OPF there is a drastic reduction in overload probabilities – the system is stable.<sup>1</sup> Moreover, this is attained with a minor increase in cost (less than five percent) while the computational time is on the order of 10 seconds.<sup>2</sup>

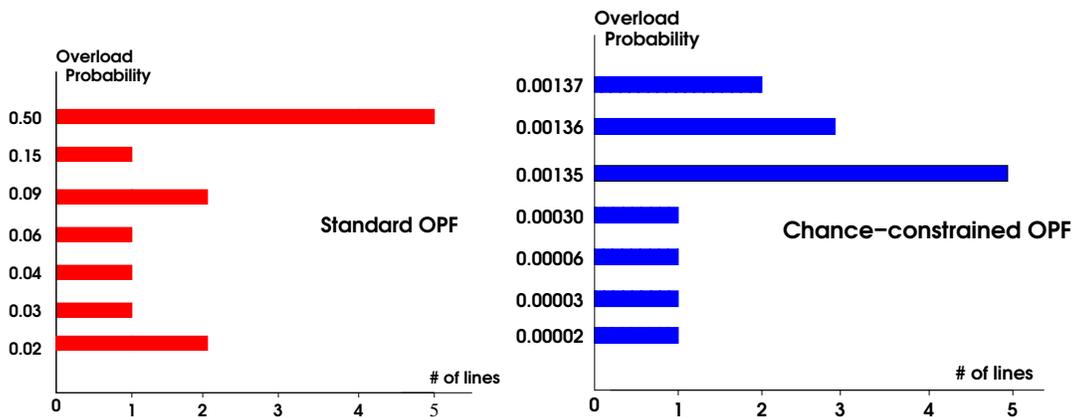


Figure 5.7: Number of lines that are overloaded with given probability values in simulation of 2746 bus Polish power grid with 20% wind penetration distributed over 50 wind farms. Under the standard OPF, five lines are overloaded half of the time, and two lines are overloaded more than ten percent of the time, constituting a situation with unacceptable systemic risk. Under the chance-constrained OPF, the largest overload probability is *fifty* times smaller than in the case of standard OPF. Moreover, the cost increase is by less than five percent.

<sup>1</sup>The chart also shows small roundoff errors produced in the run

<sup>2</sup>On a machine with two Intel Xeon CPU X5570 2.93GHz chips (total of 8 cores) and 64GB RAM.

## 5.7 Changing locations for wind farms

In this example we consider the effect of changing locations of wind farms. We take the MAT-POWER 30-bus case with line capacities scaled by .75 and add three wind farms with average power output in a ratio of 2/3/4 and standard deviations at 30% of the average. Two choices of locations are shown in Fig. 5.4. The first remains feasible for penetration up to 10% while the second can withstand up to 55% penetration. This experiment shows that choosing location of the wind farms is critical for achieving the ambitious goal of high renewable penetration.

## 5.8 Reversal of line flows

Here we consider the 9-bus case with two wind sources and 25% average penetration and standard deviations set to 30% of the average case and analyze the following two somewhat rare but still admissible wind configurations: (1) wind source (a) produces its average amount of power and source (b) three standard deviations *below* average; (2) the reverse of the case (1). This results in a substantial reversal of flow on a particular line shown in Fig. 5.8. This example suggests that when evaluating and planning for grids with high-penetration of renewables one needs to be aware of potentially fast and significant structural rearrangements of power flows. Flow reversals and other qualitative changes of power flows, which are extremely rare in the grid of today, will become significantly more frequent in the grid of tomorrow.

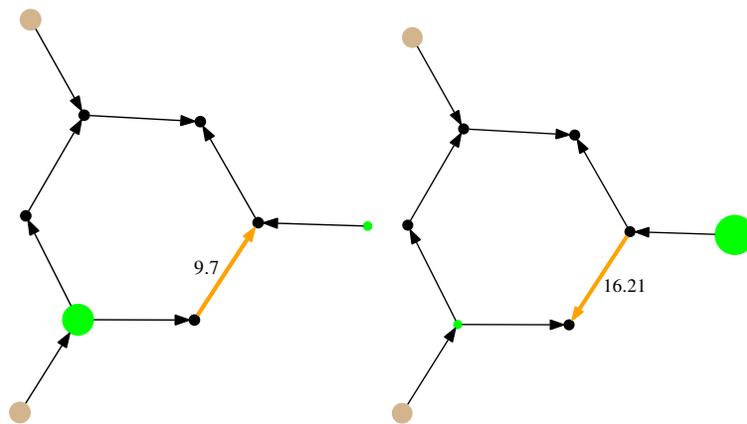


Figure 5.8: 9-bus case, 25% average penetration from two wind sources. With shifting winds, the flow on the orange line changes direction with a large absolute difference.

## **Part II**

### **The power grid attack problem**

# Chapter 6

## Introduction

### 6.1 Motivation

Large-scale blackouts of the power grid are massively disruptive to modern life. A serious blackout can leave millions without power; shut down schools, offices, and hospitals; force essential infrastructure like airports to rely on wasteful backup generators; and lead to economic losses into the billions of dollars.

As we move toward a “smart” grid, we are at the same time encountering new stresses from increased renewable generation, and growing threats from climate change and terrorism. Moreover, economic and environmental concerns have prevented the growth of power infrastructure to keep pace with rising demand, so the grid operates ever closer to capacity [42]. Because of these stresses on the modern power grid, the need to further understand the causes and properties of blackouts has never been greater.

Modern power grids are usually quite robust to small disturbances, and are typically required to operate under at least the “N - 1” criterion: the grid will continue to run safely after the failure of a single element. After the failure, grid operators act quickly to restore the N-1 status so that after a short period of time the grid can safely withstand an additional contingency. Thus, large and expensive disruptions to the grid are usually caused by the failure of several components in

a small window of time.

This motivates a common topic in the study of power grid vulnerabilities, the so-called “N - k problem” which asks: is there a set of  $k$  power lines whose simultaneous outage will result in system failure? Rather than certifying the safety of the grid, in studying this problem one hopes to identify potential weaknesses so that they might be addressed. It is an “agnostic” approach in that it does not assume any particular structure for the attack – which is useful because failures can come from a large number of sources (natural disasters, extreme weather events, physical attack, human error, etc.) Of primary interest are small values of  $k$ ; for large  $k$  (greater than 5 or so) there are usually a large number of damaging attacks, and for  $k = 1$  simple enumeration will work [20]. For large grids, combinatorial explosion makes the discrete problem intractable. Furthermore, the equations that model the physical laws of power flows need to be solved in an accurate and computationally practicable manner, which can be difficult for large grids.

### 6.1.1 Continuous formulation

One way to bypass this complexity is to consider a related continuous problem: instead of looking at power line failures, let the impedances of the lines be increased. In the limit as resistance goes to infinity, this is equivalent to removing the line. To simulate the idea of a small  $k$ , we impose constraints on how much the impedances can be changed. The hope is that this modified problem will have a combinatorial solution – the optimal attack will be to largely modify the impedance of only a small number of lines, while perhaps slightly modifying many more lines. This way, the solution to the continuous problem also answers the original discrete question. In the numerical experiments that follow we will see that this is indeed the case.

The next step is deciding what notion of “damage” to the grid we should consider. A number of factors can lead to instability in the grid and the potential for a costly blackout, as discussed in section 2.2. In [41] we investigate a variation of the problem discussed in this dissertation that seeks to maximally disrupt voltage magnitudes; unlike line overages, which usually take some time before causing the line to trip, voltage disturbances can lead to swift failures in the

grid, too fast for any mitigating action by a grid operator. Though the framework we developed can accommodate both, the examples in this dissertation focus instead on maximally disrupting voltage phase angle differences across lines. These are correlated with frequency instability and a large angle difference can lead to loss of synchrony and sudden collapse of the grid.

Our formulation is then an optimization problem<sup>1</sup>:

An attacker seeks to maximally disrupt normal bus voltages by increasing the impedances of power lines, constrained by a budget.

Note that after increasing the impedance on a number of lines, determining the voltage disruption requires recomputing the power flows with the new grid configuration, which we achieve with an OPF calculation. The approach can be described in terms of *bilevel optimization*: the lower-level problem captures the disruption to the grid arising from a given set of attacked lines, while the variables in the upper-level problem model the attack itself. Bilevel optimization can be described as “mathematical programs with optimization problems in the constraints”; [28] provides an overview of this sort of problem.

### 6.1.2 Is this realistic?

This attack model is unlikely to be encountered in practice exactly as described – it is doubtful that a malicious agent would or could raise impedances on a particular set of power lines. But it may still provide useful information about the vulnerability of the grid, as we will see in the numerical experiments.

As an example, the model often times discovers attacks that create islands, which a combinatorial formulation could be designed to find but may be too inefficient on large grids. Additionally, the structure of the attacks discovered by the model mimicks that of grid conditions immediately prior to historical blackouts: 1) a concentrated attack on a small number of lines and 2) a broad,

---

<sup>1</sup>We use the term “attacker” simply as a convenient way to describe the problem; we do not assume the grid will actually be deliberately attacked. Our method is a vulnerability analysis which can help expose weaknesses to a large variety of events which may damage the grid, including but not limited to deliberate attack.

dispersed attack over a larger set of lines which can be interpreted as systemic stress that can contribute significantly to the acute attack. So although the specific details of the attack mechanism may not be completely realistic, the model discovers sets of lines whose interdiction (even in a partial sense) lead to disruption of the grid, which should be of interest to grid planners.

## 6.2 Attacker budget and the upper-level problem

To find interesting weaknesses in the grid, we impose a budget on the attacker’s ability to increase impedances. By tuning the parameters of the budget we can find attacks of different strengths. Setting the budget too low may result in an attack which has little effect on the grid. Too high a budget on the other hand is likely to find a damaging attack, but it may be of little interest if it involves extreme changes to a large number of lines – for instance, finding an attack of “disable 25% of all power lines” is unlikely to provide useful information to grid planners and operators.

The form of the budget we use is reminiscent of the “N-k” problem discussed in the introduction to this section:

- the impedance of each line can be multiplied by a factor up to some  $\bar{\gamma}$
- the total amount of multiplication across all lines is bounded by  $\kappa\bar{\gamma}$

In this,  $\gamma$  is a vector representing the increase in branch impedances in the following way: if branch  $k$  has impedance  $z_k$ , then under the attack its impedance is

$$z'_k = (1 + \gamma_k)z_k \tag{6.1}$$

With this budget, the attacker can maximally attack  $\kappa$  lines. But unlike the “N-k” problem, the attacker may choose to allocate his budget of  $\kappa\bar{\gamma}$  across many more than  $\kappa$  lines. This sort of attack more closely resembles historical blackouts<sup>2</sup> where only a small number of lines are disabled, but many more lines are under various amounts of stress due to grid conditions.

---

<sup>2</sup>Such as the Northeast blackout of 2003 [2].

We can describe the two constraints with the following notation, from [41]:

$$\begin{aligned} e^T \gamma &\leq \kappa \bar{\gamma} \\ 0 &\leq \gamma \leq \bar{\gamma} e \end{aligned} \tag{6.2}$$

where  $e = (1, 1, \dots, 1)^T$ , the scalar  $\bar{\gamma}$  is an upper bound for each line, and  $\kappa$  is the maximum number of lines that can be attacked at the maximum level.

The upper-level problem is then to select a  $\gamma$  satisfying (6.2) which maximizes the appropriate notion of voltage disruption. In [41] we chose to maximize the sum of squared differences in voltage magnitude from 1.0 p.u.:

$$\max \sum_{k \in \mathcal{V}} (|V_k| - 1)^2$$

In this dissertation, I chose to focus mostly on maximizing the 1-norm of the vector of voltage angle differences across branches:

$$\mathcal{F}(\gamma) = \sum_{\{k,m\} \in \mathcal{L}} |\theta_k - \theta_m| \tag{6.3}$$

The choices of 1-norm versus 2-norm and angle versus magnitude cause the attacker to seek different sorts of potential weaknesses in the power grid. The power flows on the grid change after the attack via an optimal power flow computation, described in the next section. This is what produces the voltage angles  $\theta_k$  in Eq. (6.3).

### 6.3 Optimal power flow lower-level problem

After the grid is attacked, the line impedances are increased as in Eq. (6.1). We stated the OPF (see section 2.6) in terms of admittances, however, and the equivalent adjustment for the

matrices  $G$  and  $B$  is<sup>3</sup>

$$G_{km}(\delta) = \begin{cases} \frac{G_{km}}{\delta_{km} + 1} & \text{if } k \neq m, \\ -\sum_{l \in \Omega_k} G_{kl}(\delta) & \text{if } k = m. \end{cases} \quad (6.4)$$

$$B_{km}(\delta) = \begin{cases} \frac{B_{km}}{\delta_{km} + 1} & \text{if } k \neq m, \\ -\sum_{l \in \Omega_k} B_{kl}(\delta) & \text{if } k = m. \end{cases} \quad (6.5)$$

Note that when  $\gamma = 0$  the conductances and susceptances all retain their original values.

The power flow must change to account for this updated state of the grid while still satisfying the power balance equations. We accomplish this by solving an OPF satisfying these conditions:

1. generator bus voltage magnitudes remain fixed
2. generator bus power injections and voltage angles vary
3. load bus power injections remained fixed
4. load bus voltage magnitudes and angles vary
5. line flow limits are not enforced
6. the sum of squared power injection changes at the generator buses is minimized

To clarify point 6: if generator bus  $g$  has power output  $P_g^0 + jQ_g^0$  before the attack, the OPF chooses  $P_g$  and  $Q_g$  to minimize

$$\mathcal{H}(P, Q, |V|, \theta) = \sum_{g \in \mathcal{G}} (P_g - P_g^0)^2 + (Q_g - Q_g^0)^2 \quad (6.6)$$

where  $\mathcal{G}$  is the set of generator buses. Note that generator bound constraints are not enforced; the objective helps to keep generator outputs reasonable, but this is still a shortcoming that future

---

<sup>3</sup>As in the introduction to power flows in section 2.3, I have excluded the shunt terms here.

work should address. In the numerical experiments we found that strictly enforcing the generator bounds lead to many more infeasible cases, indicating that load shedding or other measures would be necessary and thus require a more complicated formulation. The OPF here is nonetheless an improvement over our approach in [41], which used a simple power flow with a slack bus. With that method, the slack bus could exhibit extremely erratic and unrealistic behavior when the grid is attacked.

This OPF represents a somewhat realistic automatic response to the disturbance introduced by  $\gamma$ : the generators adjust as little as possible to continue satisfying the demand at the load buses. Unlike the classic power flow problem with a single slack bus, this OPF distributes the change in power output across all the generators, which is a more reasonable approximation to what happens in practice.

The complete formulation of the attack problem is then:

$$\max_{\gamma} \sum_{\{k,m\} \in \mathcal{L}} |\theta_k^* - \theta_m^*| \quad \text{s.t.} \quad (6.7a)$$

$$0 \leq \gamma_{km} \leq \bar{\gamma} \quad \forall \{k, m\} \in \mathcal{E} \quad (6.7b)$$

$$\sum_{\{k,m\} \in \mathcal{L}} \gamma_{km} \leq \kappa \bar{\gamma} \quad (6.7c)$$

$$(P^*, Q^*, |V|^*, \theta^*) = \underset{P, Q, |V|, \theta}{\operatorname{argmin}} \sum_{g \in \mathcal{G}} (P_g - P_g^0)^2 + (Q_g - Q_g^0)^2 \quad \text{s.t.} \quad (6.7d)$$

$$P_k = P_k^0 \quad \forall k \in \mathcal{D} \quad (6.7e)$$

$$Q_k = Q_k^0 \quad \forall k \in \mathcal{D} \quad (6.7f)$$

$$|V_k| = |V_k|^0 \quad \forall k \in \mathcal{G} \quad (6.7g)$$

$$P_k = |V_k| \sum_{m \in K} |V_m| (G_{km}(\gamma) \cos \theta_{km} + B_{km}(\gamma) \sin \theta_{km}) \quad \forall k \in \mathcal{V} \quad (6.7h)$$

$$Q_k = |V_k| \sum_{m \in K} |V_m| (G_{km}(\gamma) \sin \theta_{km} - B_{km}(\gamma) \cos \theta_{km}) \quad \forall k \in \mathcal{V} \quad (6.7i)$$

$\delta$  is the attacker's decision variable in the upper-level problem – the amount by which to multiply the impedance of each line. Eq. (6.7a) is the upper-level (attacker's) objective, to max-

imize the 1-norm of voltage angle differences. Constraints (6.7b)-(6.7c) represent the attacker’s budget. Constraint (6.7d) states that the power flows will adjust according to the OPF in the inner-level problem expressed in (6.7d)-(6.7i). Note that constraints (6.7h)-(6.7i) differ from (2.22)-(2.23) in that the conductance and susceptance terms are now the  $\gamma$ -attacked versions  $G_{km}(\delta)$  and  $B_{km}(\delta)$  from Eqs. (6.4)-(6.5). Also observe that line flow limits are not enforced – this OPF does not include the constraints (2.20). Additionally, generator bound constraints are not enforced.

## 6.4 Related work

The topic of vulnerability assessment for power grids is of great interest and considerable research has been done on applying optimization methods to this problem. The papers [14, 53, 60] describe a bilevel approach similar to ours, using binary decision variables so that a line is either fully removed or not attacked at all, and focusing on loss of load as the attacker’s objective rather than voltage disruption. The lower-level problem in the bilevel formulation is replaced by its dual in [53] and is approximated using KKT conditions in [14].

By replacing the lower-level problem, these formulations have the disadvantage of allowing spurious solutions of the power flow equations into the model, giving a misleading picture of the effects of a prescribed attack on the grid. In this dissertation, we retain the bilevel structure, solving the lower-level problem explicitly to more reliably rule out spurious solutions. A further difference is that these papers use the linear DC approximation to describe the grid’s response to an attack rather than the more accurate AC model used in this dissertation. Because a serious attack can lead to extreme grid conditions in which the assumptions of the DC approximation do not hold, we chose to use the more complicated full-AC model.

The paper [31] presents a bilevel formulation focusing on loss of load as the attack objective, but with a lossless AC model. This is an improvement over the DC approximation, but it replaces the lower-level problem with its KKT conditions as before. They use a continuous relaxation of

the binary decision variables comparable to ours. [57] pursues a similar strategy, but retains the binary decision variables and approximates the single-level problem with a purely combinatorial reduction of the problem that bypasses direct solution of the nonlinear AC equations.

The paper [33] explores a problem closely related to ours, examining heavily congested transmission grids with infeasible configurations of generation and load. The source of the congestion is either high load or a problematic injection of uncontrollable fluctuating generation<sup>4</sup> combined with the automatic response of controllable generation. The authors explore the problem of relieving this congestion by installing a minimal set of Flexible Alternating Current Transmission System devices, which modify the reactance of the transmission lines. They use a sequential linearization method to solve the resulting bilevel optimization problem.

In [20], the problem is stated most closely to our formulation, but it uses the DC approximation: a fictitious attacker modifies reactances (rather than impedances as in the AC model) to maximally overload any power line. Because power flow on a line is a simple function of the branch angle difference in the DC model, this objective is closely related to ours.

---

<sup>4</sup>Note the similarity to part I.

# Chapter 7

## Solving the attack problem

We now describe our procedure for solving the attack problem as formulated in the previous section. To recap: we are looking for a set of impedance multipliers, constrained by a budget, that will cause maximum disruption to normal grid voltages. We formulate this as a bilevel optimization problem, where in the upper-level problem (section 6.2) an attacker selects impedance multipliers within his budget, and in the lower-level problem (section 6.3) the grid responds via an optimal power flow problem which minimizes changes in power output at the generators.

We use MATPOWER [72] to solve the OPF lower-level problem. To solve the upper-level problem, we use the Frank-Wolfe algorithm [32] for constrained optimization, which we describe in section 7.3. The problem is non-convex, so the Frank-Wolfe algorithm is not guaranteed to converge to a global optimum; we sometimes use multiple random restarts to help find a good local optimum. Our objective does not allow for a convenient analytic form for the gradient, so we use a finite-difference approximation. This is expensive on large grids, so we investigated some derivative-free approaches. We also explored a branching approach to help avoid inferior local optima and to find many different near-optimal attacks.

## 7.1 Solving the OPF lower-level problem

A primary tool in all of these analyses is MATPOWER [72], a package of MATLAB files for solving power flow problems. With MATPOWER, it is possible to use a number of internal and third-party solvers. We used the IPOPT [66] with PARDISO solver [11, 61, 62], which we observed to perform best after much experimentation. In certain cases where the OPF failed to converge, we checked with the semidefinite program solver [51] to confirm infeasibility.

## 7.2 Non-convexity and multiple local optima

The lower-level OPF subproblem is itself nonconvex and in practice may have many local solutions, as discussed in section 2.6.1. Wrapping this with an additional optimization problem does not help the issue. Standard gradient-based optimization methods (such as the Frank-Wolfe algorithm described in the next section) are thus not guaranteed to find a global optimum. We use random restarts to find a good local optimum, though this provides no guarantees either.

Fig. 7.1 provides a simple example of the problem. In this example, we attack two different lines in the 5-bus grid “case5” from MATPOWER. The objective is the 1-norm of voltage magnitude differences from 1.0 p.u.:  $\sum_{k \in \mathcal{V}} ||V_k| - 1|$ . The attack is constrained to a maximum multiplier of  $\bar{\gamma} + 1 = 3.5$  and restricted so that only lines  $\{2, 3\}$  and  $\{3, 4\}$  can be attacked. The left image shows the objective for all attacks on the two lines with  $\gamma$  values in the region  $[0, 2.5] \times [0, 2.5]$ , so that  $\kappa = 2$  (two lines can be maximally attacked). The global optimum in this case is at  $(2.5, 2.5)$ , but there is a local optimum at  $(0.66, 0)$ . Furthermore, and not shown in the figure, there is actually a local optimum at the origin, so that starting from zero results in a very weak attack.

If we set  $\kappa = 1$ , we are restricted to the region to the left and below the dotted line. In this case the global optimum is  $(0.66, 0)$  but there is a local optimum at  $(2.5, 0)$ . If we further restrict the attack to use all of the budget, so that  $\gamma_{23} + \gamma_{34} = \kappa\bar{\gamma}$ , then the feasible region is the dotted line. In this case the global optimum is at  $(.46, 2.04)$  and again  $(2.5, 0)$  is a local

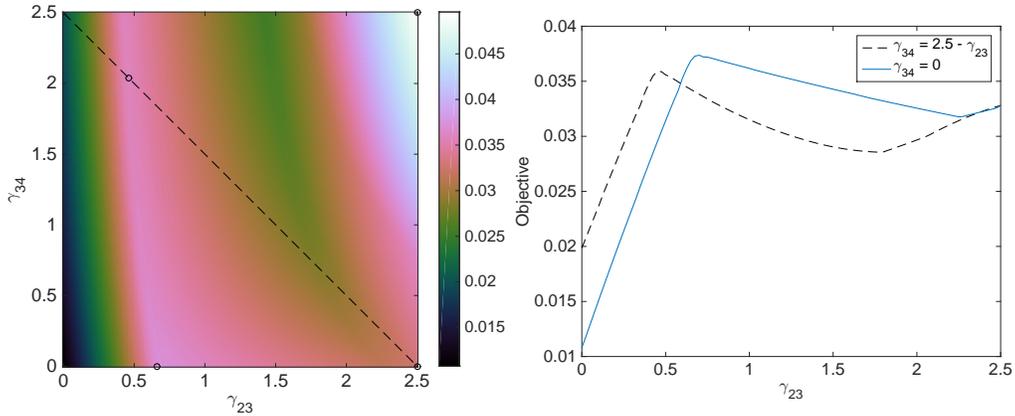


Figure 7.1: The attack problem is non-convex and can have multiple local optima. The left image shows all attacks on lines  $\{2, 3\}$  and  $\{3, 4\}$  from case5 up to a maximum of  $\bar{\gamma} = 2.5$ . The global optimum is  $(2.5, 2.5)$  but there is a local optimum at  $(.66, 0)$ . Restricting to the x-axis or the dotted line  $\gamma_{34} = 2.5 - \gamma_{23}$  results in the right image, with different local and global optima; see the text for details. Random restarts of the Frank-Wolfe procedure help to find a good local optimum.

optimum. These can be seen more clearly in the right image in Fig. 7.1.

### 7.3 The Frank-Wolfe algorithm

The Frank-Wolfe algorithm (also known as the conditional gradient algorithm) is a simple first-order method for solving constrained optimization problems. Let  $S \subset \mathbb{R}^n$  be a polyhedron and  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  a differentiable function. The Frank-Wolfe algorithm finds the solution to

$$\min_{x \in S} f(x)$$

The algorithm repeatedly solves the linearization of the problem about the current point followed by a line search, as described below:

**Algorithm 7.1:** Frank-Wolfe algorithm

Let  $k = 0$  and  $x_0 \in S$ .

**while** *not converged* **do**

1 Find a search direction  $p_k$  by minimizing the linearization of the problem at  $x_k$ :

$$y_k \leftarrow \operatorname{argmin}_{y \in S} \nabla f(x_k)^T y \quad (7.1)$$

$$p_k \leftarrow y_k - x_k$$

2 Determine a step size  $\alpha_k$  by solving (perhaps approximately) the one-dimensional optimization problem

$$\alpha_k \leftarrow \operatorname{argmin}_{\alpha \in [0,1]} f(x_k + \alpha \cdot p_k) \quad (7.2)$$

3 Update:

$$x_{k+1} \leftarrow x_k + \alpha_k \cdot p_k \quad (7.3)$$

**end**

Since  $S$  is a polyhedron, (7.1) is a linear program, and its solution  $y_k$  is an extreme point in  $S$ . Thus as  $\alpha_k$  varies from 0 to 1, the next iterate  $x_{k+1}$  varies from the current iterate to a point on the boundary of  $S$ .

We used Gurobi [10], a state of the art commercial solver, to solve the linear program (7.1). The line search in (7.2) can be solved or approximated in a number of ways; we found a standard back-tracking line search [55] to be most effective. The example in section 7.3.1 illustrates one consequence of this choice over a more exact line search. For a small  $c \in (0, 1)$  (we used  $c = .001$ ) and  $\phi \in (0, 1)$  (we used  $\phi = .5$ ), repeatedly update  $\alpha_k \leftarrow \phi \alpha_k$  (starting from  $\alpha_k = 1$ ) until the following sufficient decrease condition is met:

$$f(x_k + \alpha_k \cdot p_k) \leq f(x_k) + c \alpha_k \nabla f(x_k)^T p_k \quad (7.4)$$

For our problem, the line search typically finds a suitable step size in fewer than ten function evaluations. Especially for large grids, this is much smaller than the number of evaluations required in the finite-difference approximation to the gradient, that is to say the run-time of the line-search has little impact on the overall runtime of the algorithm.

Steps 1-3 repeat until one of the following three conditions holds:

1. improvement between iterations in the objective is less than some tolerance
2. the step size in the line search is less than some tolerance
3. the number of iterations exceeds some maximum

The third condition is a practical safeguard and we do not count it as successful convergence. Under appropriate conditions the Frank-Wolfe algorithm is guaranteed to converge to the optimal solution [24, 38], though sublinearly: convergence requires at most  $\mathcal{O}(\frac{1}{\epsilon})$  steps to get within  $\epsilon$  of the optimal value. This is typical of first-order methods – optimization methods that use only first-derivative information without taking curvature into account. As described previously, however, our problem is not well-behaved and the method does not enjoy these guarantees. But in the examples we will see that it often finds good solutions.

Though they require more iterations to converge than more sophisticated techniques, Frank-Wolfe and similar first-order methods are used widely and successfully in many fields, especially machine learning [39]. There are several reasons for this, including:

- First-order methods often have very low complexity per iteration
- In many applications, moderately accurate solutions are sufficient and can be found more quickly with a simple method
- Data is often corrupted or only known approximately, and uncertain data does not justify the precise solutions of a sophisticated method

### 7.3.1 Example

As an example of the Frank-Wolfe algorithm in action, consider the function  $f(x, y) = (x - 2)^2 + 5(y - 2)^2$  constrained to the region  $x \geq 0, y \geq 0, x + y \leq 3$ <sup>1</sup>. The gradient is  $\nabla f(x, y) = (2(x - 2), 10(y - 2))$ .

Since the feasible region is a triangle, the solution to each linear subproblem (7.1) must lie at one of the three points (0, 0), (3, 0), or (0, 3). If we start the algorithm at the origin, the gradient is  $(-4, -20)$ , and the possible solutions to (7.1) are 0, -12, and -60. Clearly the point (0, 3) minimizes the linear problem, and so the line search in (7.2) considers points along the segment from the origin to (0, 3). The exact minimum of  $f$  along that segment is at (0, 2).

If we use an exact line search, we start again at (0, 2) where the gradient is  $(-4, 0)$ . This gives solutions to the linear problem of 0, -12, and 0, and so the next search direction is towards (3, 0). The exact minimum along this direction is approximately (0.62, 1.59). With an exact line search, the solutions to the linear problem alternate in this fashion between searching towards (0, 3) and then (3, 0), producing the zig-zagging pattern as seen in the left plot in Fig. 7.2.

Using the back-tracking approach (which we've found to be most effective for the attack problem) avoids this issue, as in the plot on the right. Converging to a tolerance of  $10^{-4}$  requires 327 iterations with the exact line search, but only 7 with the back-tracking approach.

## 7.4 Computing the gradient

Since an analytic expression for the gradient isn't available for use in (7.1), we use a finite-difference approximation. Specifically, we used the standard forward difference scheme:

$$\frac{\partial f}{\partial x_i} \approx \frac{f(x + he_i) - f(x)}{h} \tag{7.5}$$

---

<sup>1</sup>The minimum is at  $(7/6, 11/6) \approx (1.17, 1.83)$ .

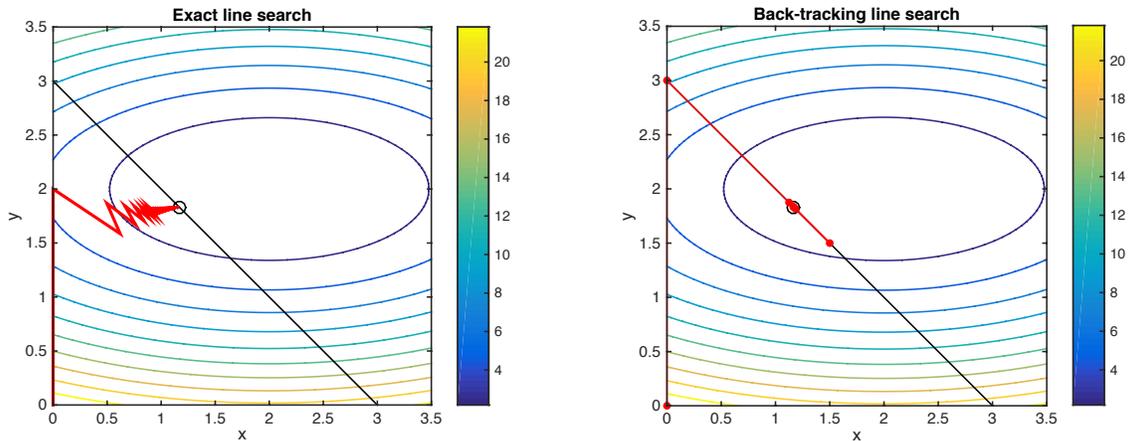


Figure 7.2: Two different line search approaches for the Frank-Wolfe algorithm applied to minimizing  $f(x, y) = (x - 2)^2 + 5(y - 2)^2$  in the region  $x \geq 0, y \geq 0, x + y \leq 3$ . An exact line search suffers from zig-zagging and takes many iterations to converge; a simple back-tracking line search quickly finds the optimum.

with  $h$  set to  $10^{-8}$  [55]. This requires  $n + 1$  function evaluations, one at  $x$  and one at each of the  $n$  perturbed points  $f(x + he_i)$ . In our case,  $n$  is the number of lines in the grid.

Computing this derivative is the major computational task of our approach, as it requires solving the inner OPF problem (section 6.3) for each line in the grid. For instance, the Polish grid example case2383wp from MATPOWER has 2896 lines, so each Frank-Wolfe iteration requires solving 2897 OPFs for the gradient computation. Because this is such an expensive process, which becomes increasingly costly as the grid size grows, derivative-free approaches appeared to be a promising alternative.

## 7.5 Other approaches

### 7.5.1 Derivative-free optimization

Because the finite-difference approximation to the gradient is expensive for large grids, we explored some derivative-free optimization methods. Pattern search or direct search [37] is a common approach for optimization problems where derivative information is not available, perhaps because the objective function is not differentiable or even continuous. It is often used in cases where the objective is an expensive black box [55], for instance the output of a simulation. So the power

Quantity	NOMAD	Frank-Wolfe
Objective	48.102	49.865
# function evaluations	129555	472
Run-time (seconds)	39996.3	52.8

Table 7.1: A typical example comparing the effectiveness of a derivative-free optimization method to our gradient-based method on the attack problem. The Frank-Wolfe method finds better attacks and requires far fewer function evaluations. Runtimes were computed from a single core of an Intel Xeon X5570 2.93GHz machine.

grid attack problem appears to be a good candidate. For certain classes of problems there are theoretical convergence guarantees, but of course our problem is not so well-behaved and these methods are not guaranteed to work well.

NOMAD [15] is a modern C++ implementation of the direct search approach. As an example, I considered a very small problem on a 30-bus grid (see subsection 8.1 for full details of the problem set-up). On this example problem, the derivative-free approach required orders of magnitude more function evaluations to discover a less-effective attack than the Frank-Wolfe approach. See table 7.1 for a more detailed comparison. The example is typical, and in fact NOMAD performed significantly worse than the gradient-based approach on every example discussed in this dissertation for which it was feasible to attempt.

## 7.5.2 Branching

Another approach to search the solution space and avoid poor locally optimal solutions is branching.<sup>2</sup> It is also a useful method to find many different near-optimal solutions – which is valuable to grid operators. It is a systematic methodology that in principle will discover any attack that can be discovered. There are many possible variations of the idea; below is a basic one:

1. Compute the best attack using a gradient-based method
2. Exclude the most-attacked line from (1) then compute the best attack

---

<sup>2</sup>Similar to the branch-and-bound method for integer programming problems.

3. Exclude the most-attacked line from (1) and 'force-include' the second most-attacked line, then compute the best attack
4. Exclude the second most-attacked line from (1) and 'force-include' the most-attacked line, then compute the best attack
5. Exclude the second most-attacked line from (1) then compute the best attack
6. ...

A concrete example of the technique on the 57-bus case is described in subsection [8.2](#). In that example, branching finds a more damaging attack than the Frank-Wolfe procedure alone, which converges to a less effective local optimum.

# Chapter 8

## Numerical experiments

In this chapter we investigate the properties of the attacks discovered by our model in a number of numerical experiments. We will begin by examining a small grid where it is easy to visualize the changes forced upon the grid by the attack. We will also look at much larger actual-scale grids, which are more difficult to visualize but we can still see that the model discovers damaging attacks. Additionally, we will consider an example of the branching procedure described in section [7.5.2](#) which can sometimes find more effective attacks by avoiding local optima.

### 8.1 9-bus example

To illustrate the nature of this attack problem, we examine a small example grid. For a small case, we can easily visualize the effect of various attacks on the grid. We can also quickly compute the result for all  $\binom{n}{k}$  attacks which allocate the entire budget across  $k$  of the grid's  $n$  lines. We will see that our procedure finds the best of these very quickly. We will also see that the impact on the grid varies widely between different attacks of the same strength.

In this example we will look at “case9” (Fig. [8.1](#)) from MATPOWER which has nine buses – three generators, three load buses with positive demand, three buses with zero demand – and nine branches.

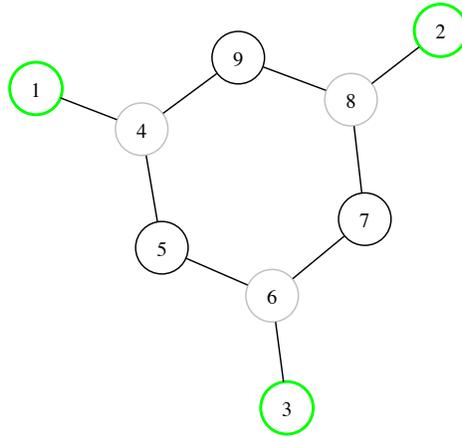


Figure 8.1: case9 from MATPOWER, rendered by graphviz [34]. Generator buses are in green, load buses are in black. Buses with neither demand nor generation are grey.

objective type	voltage angle
objective norm	$L_1$
$\bar{\gamma}$	2.5
$\kappa$	3

Table 8.1: Attack problem parameters for case9 example

### 8.1.1 Attack parameters

Let the attacker’s objective be to maximize the 1-norm of phase angle differences. We set the attacker’s budget to  $\bar{\gamma} = 2.5$  and  $\kappa = 3$ , i.e. he can multiply the impedance of a particular line by up to  $\bar{\gamma} + 1 = 3.5$  and he can maximally attack up to  $\kappa = 3$  lines.

### 8.1.2 Results

Our procedure quickly converges to allocating the entire attack budget on only three lines.<sup>1</sup> This sort of result is not typical; the optimal attack usually spreads the budget over many more than  $\kappa$  lines. But in this case it allows us to consider all of the  $\binom{9}{3} = 84$  attacks which fully allocate the budget to three lines, and see how our proposed attack compares.

<sup>1</sup>From a zero start and most random starts, it converges to this solution in two iterations

Bus	$P_g$	$Q_g$	Bus	$P_g$	$Q_g$
1	71.955	24.069	1	95.706	76.028
2	163	14.46	2	153.62	30.035
3	85	-3.649	3	79.946	13.084

(a) Unattacked grid.                      (b) Strong attack.

Table 8.2: Changes in power generation after the strong attack.  $P_g$  is real power output,  $Q_g$  is reactive. In the strong attack, lines  $\{1,4\}$ ,  $\{5,6\}$ , and  $\{8,9\}$  have their impedance multiplied by 3.5.

Our proposed attack has the greatest objective of the 84 attacks (as we would hope), and so we call it the “strong attack”. Coincidentally, the proposed attack also has the highest objective for the OPF inner problem, but this will not be the case in general. For comparison purposes, we also look at the results of the attack with the smallest objective (“weak attack”) and the median (“median attack”). See Fig. 8.2.

Though this attack focuses on the 1-norm of phase angle differences, many other measures of grid health worsen under the attack as well. The generator outputs change substantially: the lower-level OPF objective<sup>2</sup>  $\mathcal{H}$  is 3899.9, which is very large relative to the scale of the generator outputs. In particular, the reactive power output at the generators more than doubles; see table 8.2. Furthermore, the bus voltage magnitudes stray significantly from 1.0 p.u.: the 1-norm of this quantity increases from 0.101 in the unattacked grid to 0.484 under the strong attack. All this suggests that the attack discovered by this procedure is indeed a damaging attack in a broad sense. See Fig. 8.3 for more.

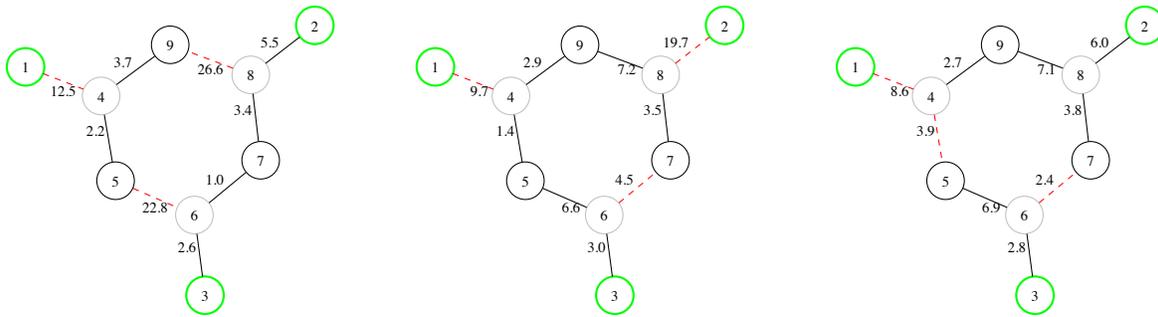
### 8.1.3 Summary

This example demonstrates three main points about this attack problem:

1. Different attacks of the same strength have greatly different effects on the grid.
2. Though the attack is optimized to increase the 1-norm of phase angle differences, it causes many other problems for the grid.

---

<sup>2</sup>Recall that this is the squared sum of differences from the generator outputs of the unattacked grid; see section 6.3



(a) Strong attack  
 $\mathcal{F} = 80.433$   
 $\mathcal{H} = 3899.9$

(b) Median attack  
 $\mathcal{F} = 58.413$   
 $\mathcal{H} = 1868.2$

(c) Weak attack  
 $\mathcal{F} = 44.129$   
 $\mathcal{H} = 106.26$

Figure 8.2: Three different attacks on case9, where a red dotted line has its impedance multiplied by 3.5.  $\mathcal{F}$  is the attacker’s objective, in this case maximizing the 1-norm of phase angle differences across branches (the units are degrees).  $\mathcal{H}$  is the objective for the OPF inner problem. The unattacked grid has  $\mathcal{F} = 33.249$  and  $\mathcal{H} = 0$ .

3. Our approach quickly finds the best attack, and we can easily verify that it is the best in this example through simple enumeration.

On larger problems, enumeration is too difficult, but our approach still finds a damaging attack reasonably quickly.

## 8.2 57-bus example: branching

In this section we provide an illustration of the branching method described in subsection 7.5.2. We consider case57 from MATPOWER with 57 buses, 7 of which are generators, and 80 lines. See Fig. 8.4 for a visualization of this grid. The attacker’s objective is to maximize the 1-norm of phase angle differences, with a budget of  $\kappa = 3$  and  $\bar{\gamma} = 2.5$ .

Starting from zero, Frank-Wolfe converges to an attack with most of the budget allocated to the lines (35, 36, 46)<sup>3</sup> with an objective of 184.83 (in degrees). If we do not allow the attack to include line 35 and restart the algorithm, it instead converges to a different attack with most

<sup>3</sup>For this and the remaining examples, we refer to line numbers rather than the  $\{k,m\}$  bus-centric notation.

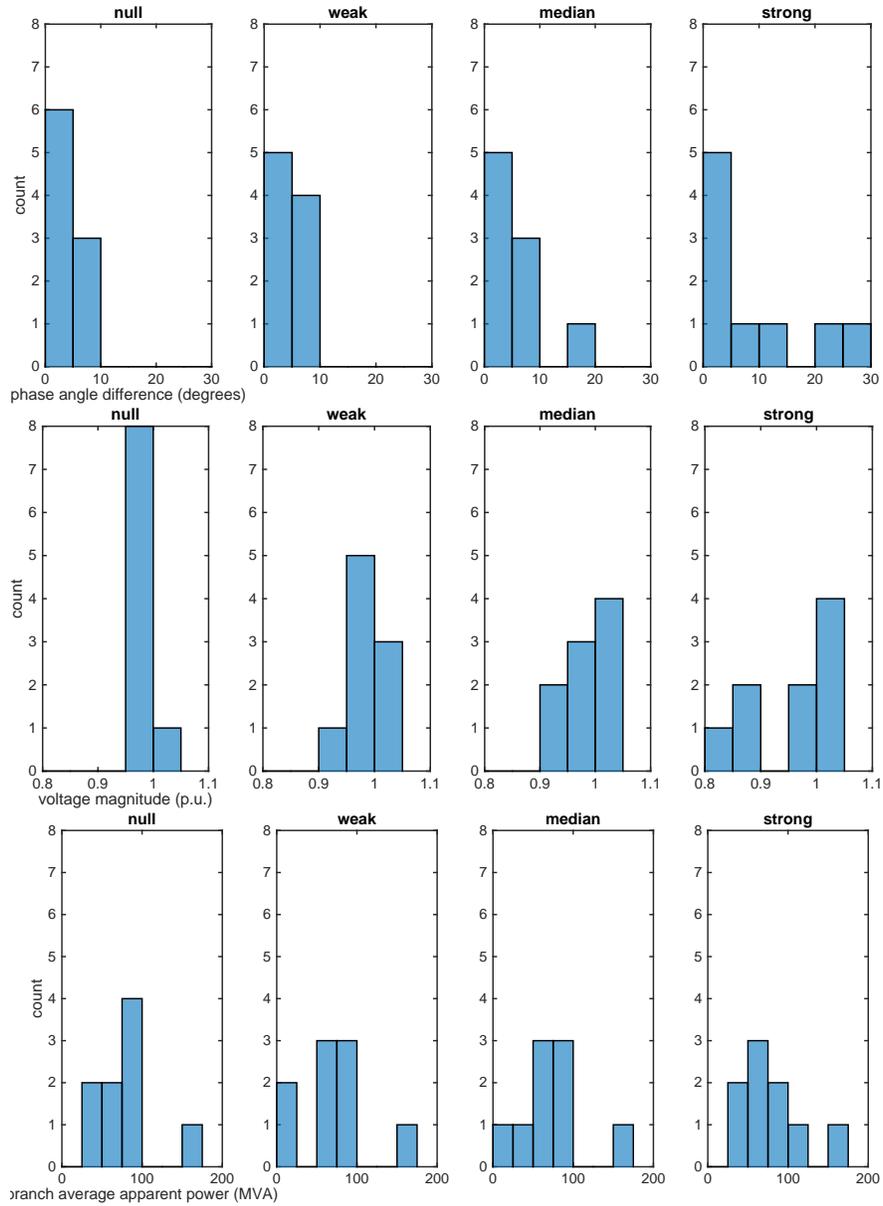


Figure 8.3: As the attack grows stronger, many different indicators of grid health grow weaker. “Null” is the unattacked grid; see the text for definitions of weak, median, and strong attacks. Branch apparent power is defined here as  $S = \frac{1}{2} \left( \sqrt{P_f^2 + Q_f^2} + \sqrt{P_t^2 + Q_t^2} \right)$  – the average of the apparent power injected into the line from the ‘to’ bus and the ‘from’ bus.

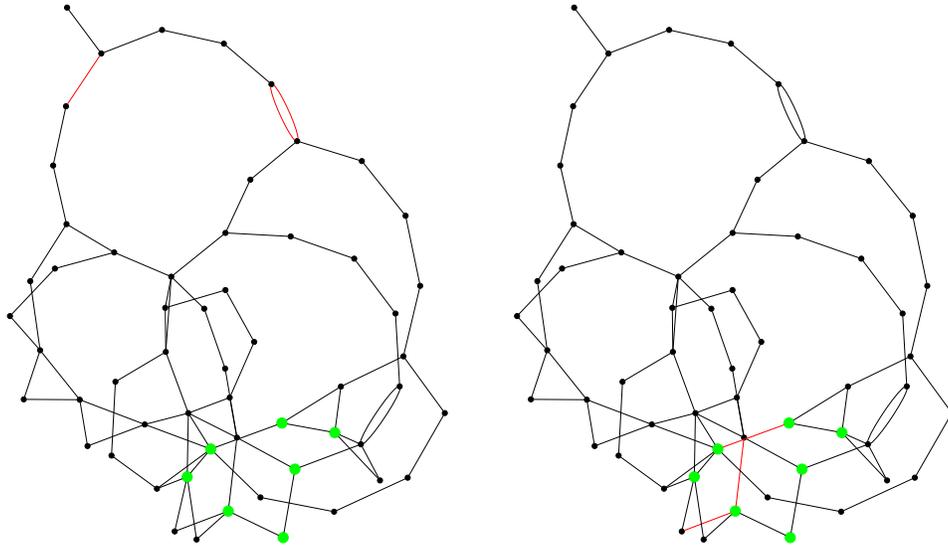


Figure 8.4: case57 from MATPOWER with generator buses in green. Pictured are two attacks discovered with the branching procedure in section 8.2. Observe that the attacks are qualitatively different: the attack on the left creates two islands, while the attack on the left focuses on generators and heavily loaded lines.

of the budget on lines (8, 15, 16), and an improved objective of 188.34. In fact, the algorithm also converges to this attack if we force-exclude lines 36 or 46. So allowing all three of these lines in the attack causes Frank-Wolfe to converge to a suboptimal local maximum, but excluding any one of them improves things. The objectives for the two different attacks are fairly close, however, so this method can also be seen as a way to find additional sets of vulnerable lines, even if the first attack found is in fact the most damaging.

An interesting observation is that the two attacks are qualitatively different: if we sever the top three lines from the first attack, it creates two islands by separating buses 25, 30, 31, 32, and 33 from the rest of the grid. The second attack focuses on lines connecting large generators to the grid, and consequently these lines are among the most heavily loaded. See Fig. 8.5 for further results from the branching procedure.

```

excluded: []
184.828 [46 1.63; 35 1.38; 36 1.35; 15 1.08; 16 1.01; 17 0.63; 2 0.29; 27 0.13; ]

excluded: [35]
188.344 [8 2.33; 15 1.71; 16 1.42; 17 1.22; 2 0.30; 18 0.29; 27 0.20; 19 0.01; 20 0.01; 25 0.00; 1 0.00; ]

excluded: [46]
188.344 [8 2.33; 15 1.71; 16 1.42; 17 1.22; 2 0.30; 18 0.29; 27 0.20; 19 0.01; 20 0.01; 25 0.00; 1 0.00; ]

excluded: [36]
188.344 [8 2.33; 15 1.71; 16 1.42; 17 1.22; 2 0.30; 18 0.29; 27 0.20; 19 0.01; 20 0.01; 25 0.00; 1 0.00; ]

excluded: [8]
184.828 [46 1.63; 35 1.38; 36 1.35; 15 1.08; 16 1.01; 17 0.63; 2 0.29; 27 0.13; ]

excluded: [8, 35]
186.926 [15 2.30; 17 1.76; 16 1.75; 2 1.55; 27 0.13; 19 0.00; 20 0.00; 25 0.00; ]

excluded: [15, 35]
179.18 [8 2.50; 17 1.73; 16 1.61; 14 0.67; 28 0.67; 36 0.15; 72 0.11; 25 0.03; 7 0.02; ]

excluded: [16, 35]
181.304 [8 2.50; 15 2.30; 2 1.26; 17 1.11; 36 0.15; 18 0.09; 66 0.05; 26 0.03; 1 0.01; ]

excluded: [8, 15, 35]
171.412 [20 2.38; 19 2.00; 16 1.16; 17 1.07; 2 0.61; 46 0.13; 36 0.05; 14 0.05; 27 0.03; 7 0.01; 54 0.00; ]

```

Figure 8.5: The branching procedure applied to case57. The first of each pair of lines incates which branches are excluded from the attack. The second is of the form: objective [branch\_number attack\_value; branch\_number attack\_value; ...]. Forcing certain lines to be excluded from the attack may find other sets of vulnerable lines, and sometimes improves the strength of the attack.

Attack	Objective
Frank-Wolfe	625
Head	612
Simple	596
Tail	471
None	464

Table 8.3: Objective values for different attacks on case118. See the text for the definitions of the different attacks.

### 8.3 118-bus example: head and tail

This example is much larger, but still small enough to visualize. The IEEE 118-bus test case (included in MATPOWER, see Fig. 8.7) has 54 generators and 186 branches. Considering all attacks of three lines at maximal strength is  $\binom{186}{3} \approx 10^6$  attacks.

But this time, as is typically the case, the optimal attack with parameters as in table 8.1 spreads the budget over many more than three lines. The majority of the budget is allocated to a small number of lines, but the remainder is spread thinly over a larger number of lines, though still a small fraction of the total number of lines in the grid. This sort of attack mimicks what has been observed in actual historical blackouts: a handful of lines are taken out, but a much larger number are stressed due to grid conditions.

We examine the relative impact of these two pieces of the attack – the strong attack on a few lines and the general “malaise” over many more lines. A typical random start of the Frank-Wolfe approach progresses as in Fig. 8.8. Notably, the majority of the budget is allocated between the six lines (38 96 31 33 67 66), the “head” of the attack. The rest of the budget is spread over a few dozen more lines, the “tail”. The unattacked grid has an objective value of 464. The Frank-Wolfe attack achieves an objective of 625. Attacking only the head and setting to zero the attack on the tail gives an objective of 612. After enumerating all possible attacks of  $\kappa = 3$  at maximal strength  $\bar{\gamma} = 5.0$ , the best result is (9 38 96) with an objective of 596. Setting the attack to zero on the head and only considering the tail gives an objective of 471. See table 8.3 for a summary and Fig. 8.6 for histograms of the phase angle differences under each attack.

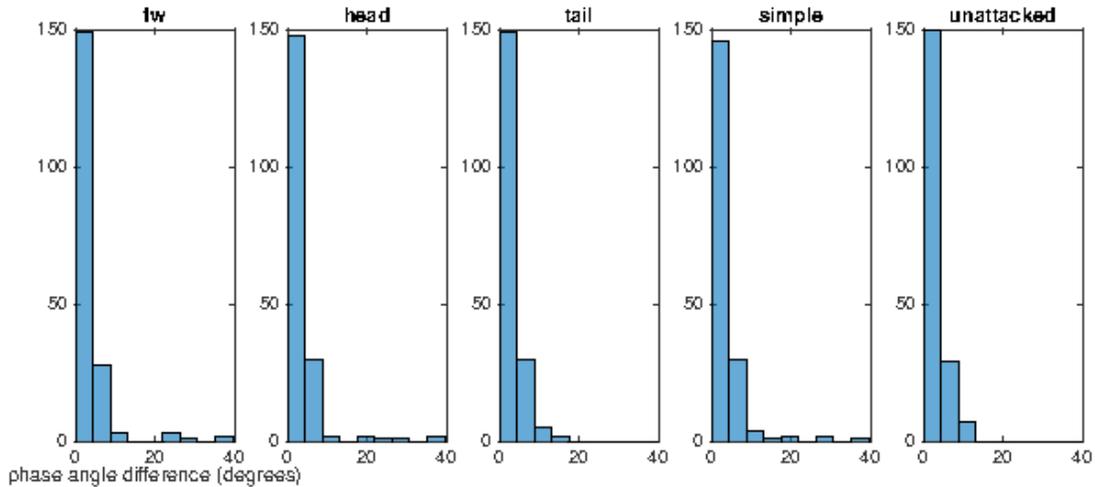


Figure 8.6: Histogram of phase angle differences under different attacks on case118. Notice that the Frank-Wolfe attack increases the phase angle difference on many branches to over 20 degrees.

This result is typical: the majority of the impact of the attack comes from hitting a few lines hard, but the weaker tail of the attack contributes significantly as well, producing a more damaging attack than focusing entirely on a few lines.

## 8.4 Large-scale example on the Polish grid

Our approach scales up to realistic, national-scale power grids. MATPOWER includes data for a number of configurations of the Polish power grid. I ran the algorithm on the 2383-bus winter 1999-2000 peak data “case2383wp.m” with 327 generators and 2896 branches.

The individual inner OPF problems are relatively expensive for problems of this size, taking up to a second to solve on a machine with two Intel Xeon CPU X5570 2.93GHz chips (total of 8 cores) and 64GB RAM. Since grids this big require solving thousands of these OPFs for each gradient, it is worthwhile to parallelize this computation. Computing the gradient is embarrassingly parallel and can be accomplished in MATLAB with a simple `parfor` loop to take advantage of all available cores.

With a budget of ( $\kappa = 5$ ,  $\bar{\gamma} = 5.0$ ) and objective to maximize the 1-norm of angle differences,

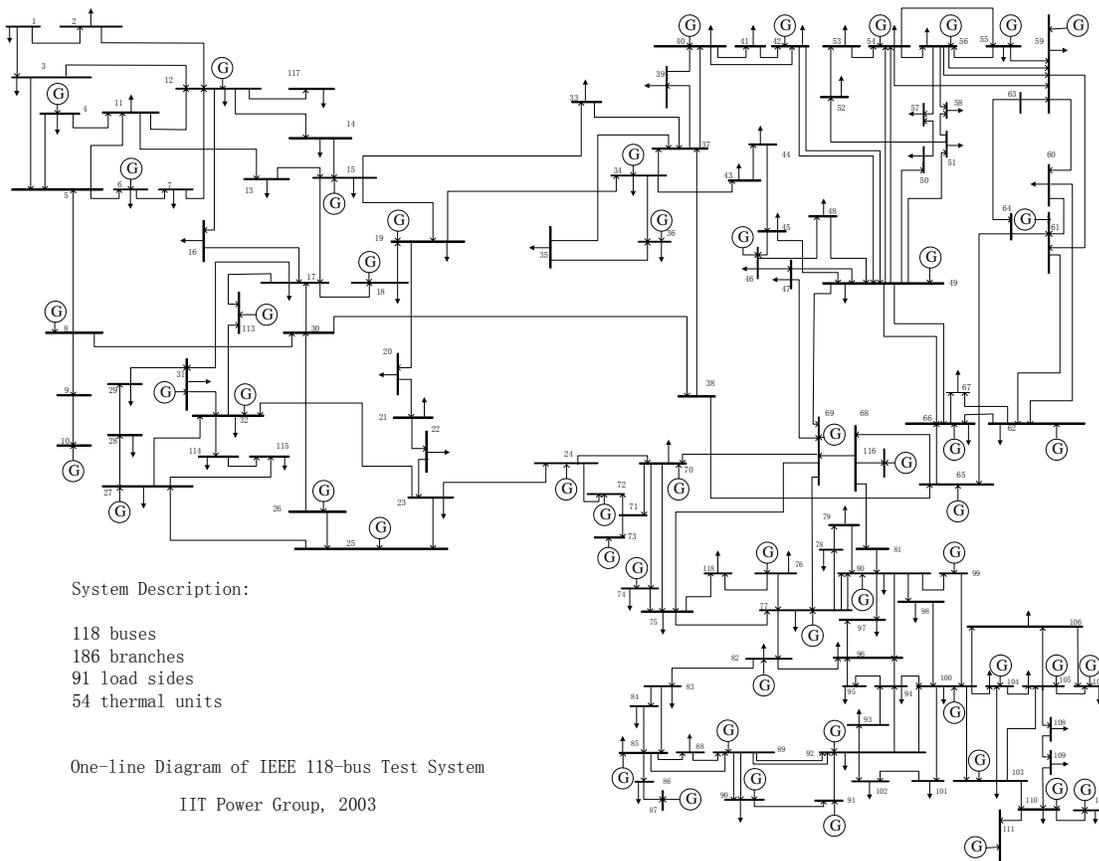


Figure 8.7: The IEEE 118-bus test system. (Image source: [1].)

```

0 obj=464.137 step=NaN [ ]
1 obj=595.511 step=1.00 [9 5.00; 38 5.00; 96 5.00; ]
2 obj=611.926 step=0.50 [9 2.50; 31 2.50; 38 2.50; 66 2.50; 67 2.50; 96 2.50; ]
3 obj=613.859 step=0.25 [96 3.12; 9 1.88; 31 1.88; 38 1.88; 66 1.88; 67 1.88; 33 1.25; 139 1.25; ]
4 obj=615.832 step=0.12 [96 2.73; 38 2.27; 9 1.64; 31 1.64; 66 1.64; 67 1.64; 33 1.09; 139 1.09; 167 0.62; ...
5 obj=616.259 step=0.25 [96 3.30; 38 2.95; 138 1.25; 9 1.23; 31 1.23; 66 1.23; 67 1.23; 33 0.82; 139 0.82; ...
6 obj=616.533 step=0.25 [38 3.46; 96 2.48; 66 2.17; 67 2.17; 138 0.94; 9 0.92; 31 0.92; 33 0.62; 139 0.62; ...
7 obj=617.755 step=0.12 [38 3.03; 96 2.79; 66 1.90; 67 1.90; 31 1.43; 138 0.82; 9 0.81; 68 0.62; 33 0.54; ...
8 obj=619.164 step=0.25 [38 3.52; 96 3.34; 33 1.65; 66 1.43; 67 1.43; 31 1.07; 138 0.62; 9 0.61; 68 0.47; ...
9 obj=619.521 step=0.06 [38 3.30; 96 3.13; 33 1.55; 66 1.34; 67 1.34; 31 1.32; 139 0.69; 138 0.58; 9 0.57; ...
10 obj=619.758 step=0.12 [38 3.51; 96 2.74; 31 1.78; 33 1.36; 66 1.17; 67 1.17; 174 0.81; 139 0.60; 138 0.5...
...
52 obj=624.790 step=0.01 [38 3.86; 96 3.11; 31 2.36; 33 1.47; 67 1.36; 66 1.36; 108 0.64; 139 0.34; 116 0.2...
53 obj=624.792 step=0.00 [38 3.86; 96 3.10; 31 2.35; 33 1.48; 67 1.35; 66 1.35; 108 0.66; 139 0.34; 116 0.2...
54 obj=624.794 step=0.01 [38 3.87; 96 3.11; 31 2.33; 33 1.51; 67 1.34; 66 1.34; 108 0.65; 139 0.34; 116 0.2...
55 obj=624.795 step=0.00 [38 3.86; 96 3.10; 31 2.34; 33 1.51; 67 1.36; 66 1.36; 108 0.65; 139 0.33; 116 0.2...
56 obj=624.796 step=0.00 [38 3.86; 96 3.10; 31 2.34; 33 1.51; 67 1.36; 66 1.35; 108 0.65; 139 0.34; 116 0.2...

```

Figure 8.8: Sample output from the Frank-Wolfe attack procedure on case118 from a zero start.

i	obj	step-size	time linesearch	time gradient
0	2555.05			
1	2642.65	1.00	2.5	538.1
2	2709.22	1.00	3.3	649.1
3	2867.52	0.50	5.4	877.2
4	2878.45	0.50	5.2	853.7
5	2880.73	0.50	5.5	847.2
6	2912.30	0.50	5.2	814.3
7	2938.54	0.25	7.1	842.7
8	2941.87	0.25	7.0	865.6
9	2966.09	0.25	7.1	878.4
10	2972.15	0.12	8.7	781.3
11	2975.11	0.06	10.2	822.1
12	2979.31	0.12	8.9	829.4
13	2984.22	0.12	8.9	835.6
14	2988.29	0.06	10.5	783.3
15	2988.37	0.01	15.1	792.9
16	2989.47	0.06	10.2	790.8
17	2989.92	0.00	17.9	827.2
18	2989.94	0.00	23.4	812.7
19	2989.94	0.00	21.6	743.9

Optimal attack:

19 obj=-2989.94 step=0.00 [90 2.37; 321 1.89; 263 1.45; 305 1.42; 309 1.42; 166 1.21; ...]

Figure 8.9: Output from the Frank-Wolfe attack procedure on case2383wp from a zero start. Included are the step-size used in the line-search step as well as run times (in seconds) for the line search and gradient steps.

the algorithm converged in 19 iterations and 4.5 hours to an attack with objective 2989.9. The per-iteration breakdown of the algorithm can be seen in Fig. 8.9. An interesting feature of the attack on this grid is that though  $\bar{\gamma} = 5.0$ , the line with most budget allocated to it has only  $\gamma_k = 2.4$ . Evidently, increasing the allocation beyond 2.4 fails to drive up the phase angle differences, and it becomes more effective for the attacker to spread his budget more thinly across many lines. A possible explanation for this is that line 90 is part of a short cycle and thus there is another route for power to flow and avoid the damaged line. The same is true for line 38 in case118 from section 8.3.

# Bibliography

- [1] *IEEE 118-bus Test System*. Electrical and Computer Engineering Department. Illinois Institute of Technology, [http://motor.ece.iit.edu/data/IEEE118bus\\_inf/IEEE118bus\\_figure.pdf](http://motor.ece.iit.edu/data/IEEE118bus_inf/IEEE118bus_figure.pdf).
- [2] *Final Report on the August 14, 2003 Blackout in the United States and Canada: Causes and Recommendations*, U.S.-Canada Power Systems Outage Task Force, 2004. <http://energy.gov/oe/downloads/blackout-2003-final-report-august-14-2003-blackout-united-states-and-canada-causes-and>.
- [3] *Integration of Renewable Resources: Transmission and Operating Issues and Recommendations for Integrating Renewable Resources on the California ISO-controlled Grid*, California ISO, 2007. <http://www.caiso.com/1ca5/1ca5a7a026270.pdf>.
- [4] *20% of Wind Energy by 2030: Increasing Wind Energy's Contribution to US Electricity Supply*, Department of Energy (DOE), 2008. <http://energy.gov/eere/wind/20-wind-energy-2030-increasing-wind-energys-contribution-us-electricity-supply>.
- [5] *Arizona-Southern California Outages on September 8, 2011*, Federal Energy Regulatory Commission and the North American Electric Reliability Corporation, 2012. <https://www.ferc.gov/legal/staff-reports/04-27-2012-ferc-nerc-report.pdf>.
- [6] *ILOG CPLEX Optimizer*, 2012. <http://www-01.ibm.com/software/integration/optimization/cplex-optimizer/>.
- [7] *MOSEK Optimizer*, 2012. <http://www.mosek.com/>.
- [8] *Wind Generation and Total Load in the BPA Balancing Authority*, 2012. <http://transmission.bpa.gov/business/operations/wind/>.
- [9] *FACT SHEET: Administration Announces New Agenda To Modernize Energy Infrastructure*, 2015. <https://www.whitehouse.gov/the-press-office/2015/04/21/fact-sheet-administration-announces-new-agenda-modernize-energy-infrastr>.
- [10] *Gurobi Optimizer*, 2015. <http://www.gurobi.com>.
- [11] A. KUZMIN, M. LUISIER, AND O. SCHENK, *Fast Methods for Computing Selected Elements of the Greens Function in Massively Parallel Nanoelectronic Device Simulations*, in Euro-Par 2013 Parallel Processing, F. Wolf, B. Mohr, and D. Mey, eds., vol. 8097 of Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2013, pp. 533–544.

- [12] M. AMIN, *Turning the Tide on Outages*, 2010. <http://www.greentechmedia.com/articles/read/turning-the-tide-on-outages>.
- [13] G. ANDERSSON, *Modelling and analysis of electric power systems*, EEH - Power Systems Laboratory, (2008).
- [14] J. M. ARROYO AND F. D. GALIANA, *On the solution of the bilevel programming formulation of the terrorist threat problem*, IEEE Transactions on Power Systems, 20 (2005), pp. 789–797.
- [15] C. AUDET, S. LE DIGABEL, AND C. TRIBES, *NOMAD User Guide*, 2009. [https://www.gerad.ca/nomad/Downloads/user\\_guide.pdf](https://www.gerad.ca/nomad/Downloads/user_guide.pdf).
- [16] S. S. BAGHSORKHI AND I. A. HISKENS, *Analysis tools for assessing the impact of wind power on weak grids*, SysCon, (2012), pp. 1–8.
- [17] A. BERGEN AND V. VITTAL, *Power Systems Analysis*, Prentice Hall, 2nd ed., 2000.
- [18] D. BIENSTOCK, *Electrical Transmission System Cascades and Vulnerability*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2015.
- [19] D. BIENSTOCK, M. CHERTKOV, AND S. HARNETT, *Chance-Constrained Optimal Power Flow: Risk-Aware Network Control under Uncertainty*, SIAM Review, 56 (2014), pp. 461–495.
- [20] D. BIENSTOCK AND A. VERMA, *The  $N - k$  problem in power grids: New models, formulations, and numerical experiments*, SIAM Journal on Optimization, 20 (2010), pp. 2352–2380.
- [21] S. BOYD AND L. VANDENBERGHE, *Convex Optimization*, Cambridge University Press, 2004.
- [22] W. A. BUKHSH, A. GROTHEY, K. I. M. MCKINNON, AND P. A. TRODDEN, *Local Solutions of the Optimal Power Flow Problem*, IEEE Transactions on Power Systems, 28 (2013), pp. 4780–4788.
- [23] M. B. CAIN, R. P. O’NEILL, AND A. CASTILLO, *History of Optimal Power Flow and Formulations*, Federal Energy Regulatory Commission. . . ., (2012), pp. 1–36.
- [24] M. D. CANON AND C. D. CULLUM, *A Tight Upper Bound on the Rate of Convergence of Frank-Wolfe Algorithm*, SIAM Journal on Control, 6 (1968).
- [25] A. CHARNES, W. W. COOPER, AND G. H. SYMONDS, *Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil*, Management Science, 4 (1958), pp. 235–263.
- [26] M. CHERTKOV, F. PAN, AND M. STEPANOV, *Predicting Failures in Power Grids: The Case of Static Overloads*, IEEE Transactions on Smart Grids, 2 (2010), p. 150.

- [27] M. CHERTKOV, M. STEPANOV, F. PAN, AND R. BALDICK, *Exact and efficient algorithm to discover extreme stochastic events in wind generation over transmission Power Grids*, CDC-ECC, (2011), pp. 2174–2180.
- [28] B. COLSON, P. MARCOTTE, AND G. SAVARD, *An overview of bilevel optimization*, Annals of Operations Research, 153 (2007), pp. 235–256.
- [29] C. DEMARCO AND T. OVERBYE, *Low voltage power flow solutions and their role in exit time based security measures for voltage collapse*, in Proceedings of the 27th IEEE Conference on Decision and Control, IEEE, 1988, pp. 2127–2131.
- [30] I. DOBSON AND H.-D. CHIANG, *Towards a Theory of Voltage Collapse in Electric Power Systems*, Systems and Control Letters, 13 (1989), pp. 253–262.
- [31] V. DONDE, V. LÓPEZ, B. LESIEUTRE, A. PINAR, C. YANG, AND J. MEZA, *Severe multiple contingency screening in electric power systems*, IEEE Transactions on Power Systems, 23 (2008), pp. 406–417.
- [32] M. FRANK AND P. WOLFE, *An algorithm for quadratic programming*, Naval Research Logistics Quarterly, 3 (1956), pp. 95–110.
- [33] V. FROLOV, S. BACKHAUS, AND M. CHERTKOV, *Reinforcing Power Grid Transmission with FACTS Devices*, arxiv:1307.1940, (2013).
- [34] E. R. GANSNER AND S. C. NORTH, *An open graph visualization system and its applications to software engineering*, Software - Practice and Experience, 30 (2000), pp. 1203–1233.
- [35] D. GOLDFARB AND F. ALIZADEH, *Second-Order Cone Programming*, Mathematical Programming, 95 (2003), pp. 3–51.
- [36] M. GRÖTSCHEL, L. LOVÁSZ, AND A. SCHRIJVER, *Geometric Algorithms and Combinatorial Optimization*, Springer-Verlag, 1998.
- [37] R. HOOKE AND T. A. JEEVES, *"Direct Search" Solution of Numerical and Statistical Problems*, Journal of the ACM, 8 (1961), pp. 212–229.
- [38] M. JAGGI, *Revisiting Frank-Wolfe: Projection-free sparse convex optimization*, 30th International Conference on Machine Learning, (2013), pp. 427–435.
- [39] A. JUDITSKY AND A. NEMIROVSKI, *First-Order Methods for Nonsmooth Convex Large-Scale Optimization, I: General Purpose Methods*, in Optimization for Machine Learning, S. Sra, S. Nowozin, and S. Wright, eds., The MIT Press, 2012, pp. 121–48.
- [40] J. KELLEY, *The Cutting-Plane Method for Solving Convex Programs*, Journal of the Society for Industrial and Applied Mathematics, 8 (1960), pp. 703–712.
- [41] T. KIM, S. J. WRIGHT, D. BIENSTOCK, AND S. HARNETT, *Vulnerability Analysis of Power Systems*, arXiv:1503.02360, (2015).

- [42] M. KLEIMAIER AND H. WEBER, *Grid Integration of Wind Generation*, in International Conference on Large High Voltage Electric Systems, CIGRE, 2009. Technical brochure 450.
- [43] A. KLOS AND J. WOJCICKA, *Physical aspects of the nonuniqueness of load flow solutions*, International Journal of Electrical Power & Energy Systems, 13 (1991), pp. 268–276.
- [44] B. KOCUK, S. S. DEY, AND X. A. SUN, *Inexactness of SDP relaxation and valid inequalities for optimal power flow*, IEEE Transactions on Power Systems, 31 (2014), pp. 642–651.
- [45] ———, *Strong SOCP Relaxations for Optimal Power Flow*, arxiv:1504.06770, (2015).
- [46] J. LAVAEI AND S. LOW, *Zero Duality Gap in Optimal Power Flow Problem*, IEEE Transactions on Power Systems, 27 (2012), pp. 92–107.
- [47] S. H. LOW, *Convex Relaxation of Optimal Power Flow. Part II: Exactness*, IEEE Transactions on Control of Network Systems, 1 (2014), pp. 1–1.
- [48] ———, *Convex Relaxation of Optimal Power Flow Part I: Formulations and Equivalence*, IEEE Transactions on Control of Network Systems, 1 (2014), pp. 15–27.
- [49] Y. V. MAKAROV, C. LOUTAN, J. MA, AND P. DE MELLO, *Operational Impacts of Wind Generation on California Power Systems*, Power Systems, IEEE Transactions on, 24 (2009), pp. 1039–1050.
- [50] L. B. MILLER AND H. WAGNER, *Chance-constrained programming with joint constraints*, Operations Research, 13 (1965), pp. 930–945.
- [51] D. MOLZAHN, *Application of Semidefinite Optimization Techniques to Problems in Electric Power Systems*, PhD thesis, University of Wisconsin-Madison, 2013.
- [52] E. C. MORGAN, M. LACKNER, R. M. VOGEL, AND L. G. BAISE, *Probability distributions for offshore wind speeds*, Energy Conversion and Management, 52 (2011), pp. 15–26.
- [53] A. L. MOTTO, J. M. ARROYO, AND F. D. GALIANA, *A mixed-integer LP procedure for the analysis of electric grid security under disruptive threat*, IEEE Transactions on Power Systems, 20 (2005), pp. 1357–1365.
- [54] A. NEMIROVSKI AND A. SHAPIRO, *Convex Approximations of Chance Constrained Programs*, SIAM Journal on Optimization, 17 (2006), pp. 969–996.
- [55] J. NOCEDAL AND S. J. WRIGHT, *Numerical Optimization*, vol. 43 of Springer Series in Operations Research, Springer, 1999.
- [56] U. A. OZTURK AND M. MAZUMDAR, *A solution to the stochastic unit commitment problem using chance constrained programming*, Power Systems, IEEE, 19 (2004), pp. 1589–1598.
- [57] A. PINAR, J. MEZA, V. DONDE, AND B. LESIEUTRE, *Optimization Strategies for the Vulnerability Analysis of the Electric Power Grid*, SIAM Journal on Optimization, (2010).

- [58] A. PRÉKOPA, *On probabilistic constrained programming*, Proc. Princeton Symposium on Math. Programming, (1970), pp. 113–138.
- [59] F. SALAM, L. NI, S. GUO, AND X. SUN, *Parallel processing for the load flow of power systems: the approach and applications*, in Proceedings of the 28th IEEE Conference on Decision and Control, IEEE, 1989, pp. 2173–2178.
- [60] J. SALMERON, K. WOOD, AND R. BALDICK, *Analysis of electric grid security under terrorist threat*, IEEE Transactions on Power Systems, 19 (2004), pp. 905–912.
- [61] O. SCHENK, M. BOLLHÖFER, AND R. A. RÖMER, *On Large-Scale Diagonalization Techniques for the Anderson Model of Localization*, SIAM Rev., 50 (2008), pp. 91–112.
- [62] O. SCHENK, A. WÄCHTER, AND M. HAGEMANN, *Matching-based preprocessing algorithms to the solution of saddle-point problems in large-scale nonconvex interior-point optimization*, Computational Optimization and Applications, 36 (2007), pp. 321–341.
- [63] J. F. STURM, *Using SeDuMi 1.02, A Matlab toolbox for optimization over symmetric cones*, Optimization Methods and Software, 11 (1999), pp. 625–653.
- [64] R. H. TÜTÜNCÜ, K. C. TOH, AND M. J. TODD, *Solving semidefinite-quadratic-linear programs using SDPT3*, Mathematical Programming, Series B, 95 (2003), pp. 189–217.
- [65] M. VRAKOPOULOU, K. MARGELLOS, J. LYGEROS, AND G. ANDERSSON, *A probabilistic framework for security constrained reserve scheduling of networks with wind power generation*, in Energy Conference and Exhibition (ENERGYCON), 2012 IEEE International, 2012, pp. 452–457.
- [66] A. WÄCHTER AND L. T. BIEGLER, *On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming*, Mathematical Programming, 106 (2006), pp. 25–57.
- [67] G. WANG, M. NEGRETE-PINCETIC, A. KOWLI, E. SHAFIEEPOORFARD, S. MEYN, AND U. SHANBHAG, *Dynamic Competitive Equilibria in Electricity Markets*, in Control and Optimization Theory for Electric Smart Grids, A. Chakraborty and M. Illic, eds., Springer, 2011.
- [68] H. WANG, C. E. MURILLO-SANCHEZ, R. D. ZIMMERMAN, AND R. J. THOMAS, *On Computational Issues of Market-Based Optimal Power Flow*, IEEE Transactions on Power Systems, 22 (2007), pp. 1185–1193.
- [69] J. WANG, M. SHAHIDEHPOUR, AND Z. LI, *Security-Constrained Unit Commitment With Volatile Wind Power Generation*, Power Systems, IEEE Transactions on, 23 (2008), pp. 1319–1327.
- [70] H. ZHANG, *Chance Constrained Programming for Optimal Power Flow Under Uncertainty*, Power Systems, IEEE Transactions on, 26 (2011), pp. 2417–2424.

- [71] S. ZHANG, Y. SONG, Z. HU, AND L. YAO, *Robust optimization method based on scenario analysis for unit commitment considering wind uncertainties*, in Power and Energy Society General Meeting, 2011 IEEE, jul 2011, pp. 1–7.
- [72] R. ZIMMERMAN, C. MURILLO SANCHEZ, AND R. THOMAS, *MATPOWER: Steady-State Operations, Planning, and Analysis Tools for Power Systems Research and Education*, Power Systems, IEEE Transactions on, 26 (2011), pp. 12–19.