Title
Optimization Tools for Constrained Energy Markets

Author
Munsing, Eric

Publication Date
2018

Peer reviewed|Thesis/dissertation

Optimization Tools for Constrained Energy Markets

by

Eric Munsing

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

in

Engineering - Civil and Environmental Engineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Assistant Professor Scott J. Moura, Chair
Professor Duncan S. Callaway
Professor Raja Sengupta

Summer 2018

Optimization Tools for Constrained Energy Markets

Abstract

Optimization Tools for Constrained Energy Markets

by

Eric Munsing

Doctor of Philosophy in Engineering - Civil and Environmental Engineering

University of California, Berkeley

Assistant Professor Scott J. Moura, Chair

This dissertation develops an interdisciplinary approach to integrating renewable energy resources into energy markets, using tools from optimization theory, power systems, and economics. It advances prior work with the development of a set of tools for securing distributed and fully-decentralized optimization problems, including both algorithmic guards against attacks by malicious nodes, and system architectures which can enable decentralized security checks. Leveraging the emerging technologies of *blockchains* and *smart contracts*, it develops a new paradigm of blockchain-secured distributed optimization, demonstrated with simulation of a microgrid which is able to securely operate without oversight from a utility or central operator.

As the goal of interdisciplinary research is to show mastery of each field and extend current knowledge by exploring their intersection, the chapters of this dissertation are designed to support that goal:

- Chapter 1 provides background on the technical challenges of integrating high amounts of renewable energy into the electricity system, and discusses technologies and policies which can address those challenges.
- Chapter 2 introduces the idea of applying optimization tools to energy systems by studying a small-scale energy harvesting system in which batteries and capacitors are used to meet a defined load, and constraints force the use of nonlinear optimization techniques.
- Chapter 3 explores how large-scale energy storage systems can be designed and sited to maximize profits from participating in wholesale energy markets, using a linear program which demonstrates how convex optimization tools can be united with energy market data to create scalable tools for modeling large networks.
- Chapter 4 expands the study of market-based models by examining the strategic operation of generation resources on a congested network. We use game theory to model participants' behavior, power flow models to reflect the underlying constraints of the

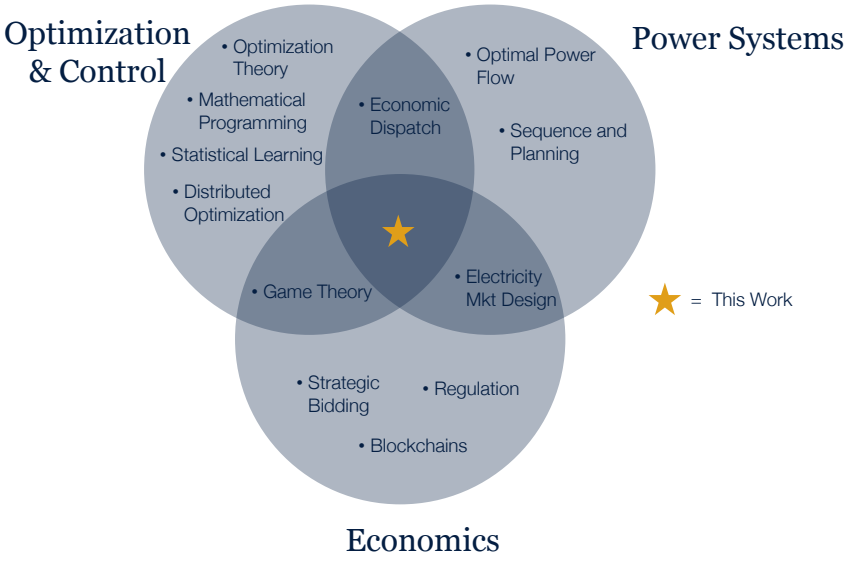


Figure 0.1: Overview of the fields which are used in the current research, showing the dissertation's location at the intersection of optimization tools, power systems engineering, and economics.

physical network, and robust convex optimization to explore how uncertainty is integrated into decision-making.

- Chapter 5 discusses how decentralized optimization models can facilitate scalable optimization tools, and explores the security risks which these optimization models introduce. Tools for detection and mitigation of attacks are explored and tested on a simple problem. Potential architectures for securing decentralized optimization are explored, including *blockchains* and *smart contracts*.

- Chapter 6 extends this approach by developing a blockchain-secured fully-decentralized optimization for a microgrid dispatch problem, coordinating distributed energy resources. By demonstrating the usefulness of this blockchain-secured optimization model, this culminating chapter shows how difficult optimization models can be scalably solved in a manner which respects privacy while guaranteeing security.

- The Appendices present tutorial material on the tools used throughout the paper, and are supplemented by the code in the author's github repository: `https://github.com/emunsing/tutorials`

Each of the chapters is independent and self-contained, and a reader familiar with the underlying tools is encouraged to jump to the chapter of greatest interest. Figure 0.1 depicts how the different topics which are addressed in this dissertation fall into conventional academic fields, illustrating how the current work draws from each.

The following associates each of these topics with specific chapters, and provides a canonical reference which can provide the unfamiliar reader with the necessary background in the topic:

- Optimization Tools
  - Nonlinear Optimization & Particle Swarm Optimization ([1], Chapter 2)
  - Convex Optimization ([2], Chapters 3, 4, 5, 6)
  - Robust Optimization ([3], Chapter 4)
  - Decentralized optimization and the Alternating Direction Method of Multipliers ([4], Chapters 5,6)
  - Fully-decentralized optimization ([5], Chapters 5,6)
- Economics
  - Microeconomics ([6] Chapter 3, 4)
  - Game theory and strategic bidding ([7] Chapter 4)
  - Blockchains and smart contracts ([8], Chapters 5, 6)
- Power Systems
  - AC and DC power flow models ([9], Chapter 4)
  - Distflow and local power flow models ([10], Chapter 6)

This dissertation is intended to serve as a springboard for future researchers studying smart grid systems, particularly those who are interested in developing secured fully-decentralized optimization models which address constrained systems.

To My Wife, Galina

Who trusted in me even when I could not.

Contents

Contents **ii**

List of Figures **iv**

List of Tables **ix**

1 Background and Motivation **1**
1.1 Motivation 1
1.2 Background: High Renewable Energy Penetration 2
1.3 Potential Solutions 7
1.4 Costs 11
1.5 Concluding Notes 11

2 Design of Small-Scale Storage Systems **14**
2.1 Introduction 14
2.2 System Model 15
2.3 System Design 16
2.4 Results 17
2.5 Conclusions 18
2.6 Potential Improvements 18

3 Transmission-Scale Storage Sizing and Siting **20**
3.1 Introduction 20
3.2 Modeling Assumptions 22
3.3 Formulation 23
3.4 Results 26
3.5 Discussion 35
3.6 Potential Improvements 37

4 Strategic Equilibria in Congested Energy Markets **39**
4.1 Introduction 39
4.2 Problem Formulation 42
4.3 Example and Results 49

4.4 Conclusion . . . 54
4.5 Potential Improvements . . . 54

5 Fully-Decentralized Optimization and Security 56
5.1 Motivation . . . 56
5.2 Background Literature . . . 57
5.3 Outline . . . 60
5.4 Mathematical Background and Notation . . . 61
5.5 Attack Vectors in Decentralized Optimization . . . 64
5.6 Developing a Detection Algorithm for Noise-Injection Attacks . . . 67
5.7 Simulation and Results . . . 73
5.8 Limitations . . . 77
5.9 Extensions . . . 78
5.10 Extension: Fully-decentralized optimization . . . 79
5.11 Attack Vectors in Fully-Decentralized Optimization . . . 80
5.12 Summary of Security Challenges . . . 83
5.13 Potential Improvements . . . 86

6 Blockchains and Energy Control 87
6.1 Introduction and Motivation . . . 87
6.2 Blockchains and smart contracts . . . 89
6.3 Prior Literature . . . 90
6.4 Optimal Dispatch Formulation . . . 93
6.5 Blockchains and ADMM . . . 99
6.6 Implementation: Test network . . . 101
6.7 Limitations . . . 106
6.8 Conclusions . . . 106
6.9 Potential Improvements . . . 107

7 Conclusion 108

Bibliography 109

A Decentralized Security- Analytic Solution 124
A.1 Analytic Solution Notes . . . 124

List of Figures

0.1 Overview of the fields which are used in the current research, showing the dissertation's location at the intersection of optimization tools, power systems engineering, and economics. 2

1.1 Forecasts for a "duck curve" with high renewable penetrations. Source: CAISO [11] 3
1.2 Levelized cost of flexibility for transmission, CCGT, demand side solutions, and storage. Costs from [12] and figure from [13] 12
1.3 Summary of Flexibility Resources 12

2.1 Equivalent Circuit Model; the design can be adjusted by changing the number of TEG elements in series in each string, or adding parallel TEG strings, supercapacitors, or batteries in parallel. 15
2.2 Radio power draw, with inset detail of the data acquisition/transmission period. 15
2.3 Bus voltage during radio dispatch, showing how the nonlinear optimization reduces costs by operating closer to system bounds. 18

3.1 Example of Location Marginal Price (LMP) distribution on the CAISO grid. Each circle represents an LMP node on the the CAISO grid. Data from 4PM PDT, August 18 2013 25
3.2 Example of the impact of reservoir size h on optimal charging behavior, while system efficiency is held constant. The smaller 2-hour system is able to more quickly charge and discharge, creating larger average price differences for each transaction (and thus greater profits per average cycle). The larger reservoir allows the system to take advantage of sustained price differences- increasing net profits but reducing average profits per cycle. The larger system also incurs greater construction costs, and so optimal system design thus balances reservoir size with construction costs. 27
3.3 Example of the impact of storage system round-trip efficiency on optimal charging behavior, assuming a constant reservoir size. The storage system will only discharge when price differences are greater than the cost of efficiency losses, leading the high-efficiency system to cycle more often. 28

3.4 Example of the impact of reservoir sizing on long-run operator profits (including the cost of constructing the reservoir). At low reservoir sizes, the operator is not able to take advantage of all arbitrage opportunities. At large reservoir sizes, construction costs outweigh arbitrage benefits. At optimum, the marginal benefit of additional reservoir capacity is balanced by construction costs. Data is shown for the BARRY_6_N001 node in 2013 and assumes an amortized reservoir cost \$5/kWh/year. 29
3.5 Optimal reservoir size in hours of storage capacity with varying annualized construction costs γ, across all CAISO LMP nodes. Optimal size of the storage system decreases when construction costs increase, as the marginal benefits from a larger reservoir are more rapidly offset by construction costs. 30
3.6 Short-run arbitrage profits for each node, plotted at varying efficiencies for a 1kW/1kWh system. Each node is represented by a line plotted at 1% opacity to show the distribution of values at each efficiency. 31
3.7 Annual cycle count at each node with varying efficiency, 1kW/1kWh system. Thresholds for daily and twice-daily charging are shown, and highlight the significance of diurnal price patterns. Each node is represented by a line plotted at 1% opacity to show the distribution of values at each efficiency. 32
3.8 Profit per cycle for each node under varying efficiencies, assuming a 1kW/1kWh system. Each node is plotted at 1% opacity to show the distribution of values. . 32
3.9 Impact of price responsiveness on storage operator profits. We assume that the market price of energy is depressed by a factor α whenever the storage operator sells energy, and is increased by α when the storage system buys energy. While profits decrease, the distribution of the profits across nodes is preserved. 34
3.10 Histogram of nodal profits at 90% efficiency for 1kW/1kWh system, with best fit of a normal distribution (estimated with maximum likelihood estimation) 34
3.11 Plot of short-run trading profits from energy arbitrage of a 90% efficient 1kW/1kWh system at LMP nodes on the California ISO grid. Broad regions of high-value nodes in the Northern and Southern coastal regions suggest large regions of congestion where storage may have greatest impact. 36

4.1 3 Node Network 49
4.2 Profits of both generation firms under a range of uncertainty intervals. Both robust and non-robust strategies are shown, and the three lines indicate high, expected, and low realizations of demand. The robust equilibrium increases prices for low uncertainty intervals, and limits exposure to downside risk. 50
4.3 Profits of each generating firm when congestion is present on the network. Note that the shape of the curves changes over distinct domains, dictated by the congestion on the network: in domain a line 1-3 is congested; in domain b lines 1-3 and 1-2 are congested, and in domain c only line 1-2 is congested. 51

4.4 Consumer surplus under a number of equilibrium models: perfect competition, Nash-Cournot equilibrium, and Nash-Cournot robust equilibrium. When producers restrict production to be robust to uncertainty, consumers are clearly impacted. Congestion further reduces consumer surplus by introducing congestion charges. 53
4.5 Net Social Benefit (sum of consumer surplus, producer profits, and merchandising surplus) under both robust and non-robust equilibrium. 53

5.1 Different structures for solving mathematical optimization problems, from left: Centralized optimization, where all details of objective and constraints are held by a central entity. Decentralized optimization (also called aggregator-coordinated optimization) where local nodes hold local objective and constraint information, and an aggregator brings nodes into consensus on shared constraints. Fully-decentralized optimization, where no centralized entity exists but neighbors communicate directly with each other to achieve consensus. 57
5.2 Publication frequency for the topics of 'Distributed Optimization', 'Cybersecurity', and the intersection of the two fields, 2010-2017. Legend captions include the total number of publications over this period; the number of publications at the intersection is three orders of magnitude less than each field. 58
5.3 Graphical example of how an attacker may distort the constraint set from $\mathcal{X}$ to $\tilde{\mathcal{X}}$ to create an optimum outside of the truly feasible set, and the limited ability to mitigate these impacts by projecting onto a publicly knowable constraint set $\mathcal{X}_{\text{pub}}$. 66
5.4 Plot of local residuals in a 3-node system under noise-injection attack. Injection of noise prevents problem convergence at nodes across the network. 66
5.5 Conceptualization of the attack detection process for a $n = 2$ toy problem. Although the validator cannot directly assess the private objective function $f(x_1, x_2)$ (shown in red-blue gradient), they are provided with a sequence of iterates x^k, and can to use a subset of these for the detection algorithm, shown here as (x_1, x_2) points with $f(x) = 0$ due to the unknown objective value. Using information from the sequence of iterates, the validator can then assess the implied gradient of $f(x)$ at each of these points. Finally, from these gradients, the validator can estimate the Hessian, and use this to evaluate the convexity of the (unknown) objective function. This can be conceptualized as a local quadratic approximation of the objective function at the reference point, shown in blue-green. 68
5.6 Visualization of the relationship between the iterates x^i, x^j, the function surface $f(x)$, and the gradient evaluations $\partial f(x)|^i, \partial f(x)|^j$. While the function surface $f(x)$ and its Hessian cannot be directly evaluated, the change in gradient evaluations between x^i and x^j can be used to derive an approximation of the Hessian. 69

5.7 The introduction of noise shifts the gradients computed using the above algorithm, shown here as shifting the unattacked (gray) vector to a new (attacked, red) position. It can be readily seen that a sufficient displacement will cause the estimated curvature (Hessian) to become nonconvex, as shown by the inferred surface represented in Subfigure (b). To avoid detection, an attacker would thus need to use a very small injected signal, but this would not be sufficient to stop convergence and the attack would fail. 72
5.8 Depiction of misclassification errors in the noise detection algorithm. In weakly convex regions of a function, numeric conditioning errors can lead to false positives (subfigure a), whereas problems which have a very strong gradient may converge before the noise detection algorithm is able to identify the injection of noise into the algorithm (subfigure b). In both cases, the algorithm for selecting points used to assess the gradient can lead to more accurate assessment of the Hessian, reducing misclassification errors. 73
5.9 Convergence results for 10,000 QPs, simulated both with and without noise-injection attack. Without attack, all simulations converge in less than 300 iterations. Under attack, 86% of problems do not converge in 500 iterations (at which time calculation was stopped). 76
5.10 Fully-decentralized optimization problem structure, highlighting two nodes and noting how this can be indefinitely expanded with the addition of further upstream and downstream nodes. Each node holds private information on its own objective function, constraints, and dual variable (e.g. u_x^k), and accepts updates from its immediate neighbors. 80
5.11 Potential attack scenarios in a fully-decentralized optimization scenario, where the z-update node is attempting to establish the veracity of a received update $x^{?k}$. Without knowing details of the private objective functions $f(x), h(w)$ and constraints $x \in \mathcal{X}, w \in \mathcal{W}$ it is difficult to detect and localize (or mitigate) an attack. The z-update node is not generally able to determine the difference between the unattacked scenario shown in (a), an attack by the immediate upstream neighbor $\tilde{x}^k$ as in (b), and an attack by a node further upstream such as $\tilde{w}^k$ as shown in (c). 81
5.12 Potential architectures for allowing security checks for fully-decentralized optimization algorithms. 85
5.13 Illustration of how a blockchain-based network provides comparable security to an aggregator-coordinated architecture 85

6.1 Sample problem structure, with arrows showing shared variables. Additional connections to upstream and downstream nodes may also be considered. The node computing $g(z), z \in \mathcal{Z}$ has been compromised. 88

6.2 Symbolic representation of the data in a blockchain, showing blocks B^0 to B^{t+1} with detail of block B^t. Blocks are linked by their cryptographic hashes $\Upsilon(B^t)$, securing the contents from alteration and allowing transparent auditing of system history. Messages M_i^t contain information about changes to the system state, such as energy transfers or payments. 89
6.3 Comparison of an efficient market in which the optimal quantity $Q*$ is cleared at price $P(Q*)$, and a market operated by a monopoly who is able to charge separate prices for generation and consumption. In this model, the monopoly restricts output to Q_M, purchasing energy at $C(Q_M)$ and charging consumers $P(Q_M)$. 91
6.4 Conceptual models of the decentralized optimization involved in blockchain consensus networks, decentralized optimization problems, and our (novel) combined paradigm. 92
6.5 The 55-bus sample microgrid test feeder used in the simulation, with a microturbine placed at Bus 1 and DERs randomly distributed throughout the network. . 101
6.6 Schedule of commitments generated by the ADMM algorithm and stored to the smart contract. Positive values of power indicate real power consumption, and negative values indicate generation/injections. 102
6.7 Voltage magnitude at each of the buses on the test network, for each hour in the simulation. Voltages vary based on local injections, and variations in time can be seen due to the impacts of local DER scheduling. 103
6.8 Convergence of the fully-decentralized ADMM algorithm under 3 modes of operation: normal operation (no compromised nodes), under the presence of attack from a node which is conducting a convergence-stalling attack, and under attack but secured with blockchain-based security checks. 104
6.9 Scheduled voltage magnitude at each of the buses on the test network at each hour in the simulation, under three operating scenarios: normal operation (no nodes attacked), attacked by a malicious node which solves a privately infeasible problem, and a blockchain-secured algorithm. The security checks enabled by the blockchain system allow the network to reach a safe equilibrium in the presence of attack. 105

List of Tables

2.1 System designs resulting from conventional engineering and optimization-based approaches; a 55% decrease in cost is shown. 18

5.1 Results for a simulation of attack detection. 500 quadratic programs were generated without attack, and false positives (upper right quadrant) were identified. An additional 500 quadratic programs were generated with simulated attack, and false negatives measured. 76

5.2 Security Issues in aggregator-coordinated and fully-decentralized systems 84

Acknowledgments

This has been interdisciplinary work, and I have benefited from the many faculty, postdocs, and students with whom I've had the pleasure to collaborate. I specifically would like to acknowledge:

- Jonathan Mather in the lab of Kameshwar Poola has been an incredible partner-in-crime on a variety of research projects,
- Martin Cowell, in the lab of Paul Wright, has been a great support in answering questions in both research and life,
- Caroline LeFloch, my predecessor in the eCal Lab and role model in entrepreneurship,
- Bertrand Travacca, whose wit is as quick as his matrix transformations and his adventures are as bold as his conjectures, and
- Laurel Dunn, whose scientific mindset provides the tie between our engineering fantasies and real-world data.

The eCAL lab is still in its infancy, and it has been great to know each of the students, postdoctoral scholars, and students who have passed through its doors. I appreciate the work of Hector Perez, Eric Burger, and Caroline LeFloch in pioneering a lab culture that values rigor, impact, and real-world applicability above all else. To Laurel Dunn, Bertrand Travacca, Saehong Park, and Sangjae Bae- the lab is now yours to shape.

This degree has challenged me more than anything else I have ever undertaken. Some remarkable faculty have supported me along the way, even when I was not ready to rise to the challenge they presented: Laurent El Ghaoui, Severin Borenstein, and Claire Tomlin, your classes forced me to grow when I was not ready for it.

I would also like to thank Professor Duncan Callaway and Professor Raja Sengupta for being part of my dissertation committee, and giving me feedback on this work.

Berkeley has been a great community over the five years that I have spent here, and I have benefited enormously from the connections I have made outside of classes:

- **BERC**, the Berkeley Energy and Resources Collaborative, which has provided an incredible community of energy geeks who are excited to discuss the latest policy and technology developments,
- **CHAOS**, the Cal Hiking and Outdoors Society, which kept me grounded and connected with the outdoors rather than stuck in a lab,
- **Blockchain at Berkeley**, an incredibly enthusiastic group of students creating a brave new world.

The financial support necessary to complete my doctoral education and this work was made possible by the National Science Foundation Graduate Research Fellowship Program, the Graduate Division, the Civil and Environmental Engineering Department, and the Energy, Controls, and Applications Lab. I thank all of these sources for allowing me to focus on my doctoral studies and research throughout my stay.

Whether spending a month on a glacier, climbing to the summit of El Capitan, or counseling me in the dark moments of the dissertation, my wonderful wife Galina Melamed has been my endless cheerleader at every step of the way.

Finally, I would like to thank my parents Beverly and Peter Munsing for their endless love and support throughout my doctoral studies. Without them I would not be the person I am today.

Chapter 1

Background and Motivation

1.1 Motivation

The production of electricity from wind and solar energy has become a key element of climate change mitigation strategies, and 29 US states currently have set procurement targets through Renewable Portfolio Standards [14]. Some of these are quite aggressive- California is considering a 50% RPS by 2030 [15], and Hawaii recently legislated a 100% RPS by 2045.

These renewable portfolio standards may be essential to meeting aggressive greenhouse gas (GHG) emissions reduction targets both for the electricity grid and the wider economy [16, 17]. However, the renewable energy sources which they promote are typically characterized by intermittent and variable generation patterns, which may lead to significant curtailment of renewable generation during peak hours and to increased reliance on natural gas peaker plants to supply ramping response [15] - two outcomes that are environmentally and economically undesirable.

We survey a number of options that could help the electricity grid adapt to increased renewable energy integration in line with high Renewable Portfolio Standards. These can broadly be described as contributing to the 'flexibility' of the grid, enabling it to respond to the variability and intermittency inherent in renewable generation systems [12].

To date, much of the grid's demand-side flexibility has been provided by demand response programs, and this could be expanded in the future to address increased renewables penetration [18]. We discuss how current demand response approaches might be expanded to integrate automation systems and controllable loads that can be scheduled in advance.

The storage of electricity has been highlighted as a key technology with considerable potential for addressing renewable energy integration [19], and could add considerable flexibility to both transmission and distribution networks. To promote the development of storage systems, the California Public Utility Commission has mandated that Californian utilities secure 1.3 GW of electricity storage systems by 2020, 700 MW of which would be at the transmission level [20], and other regions are considering similar policies.

Models of future carbon-constrained grid scenarios have shown the key roles that these

technologies may play [21–23], but highlights that the ultimate mix of flexibility will be strongly subject to system costs and carbon policies.

We briefly motivate this study with a survey of the key opportunities and challenges presented by the drive towards a higher renewable portfolio standard, and explore the technical and market challenges which renewable resources present.

1.2 Background: High Renewable Energy Penetration

Opportunities

The electricity sector accounts for approximately 40% of US greenhouse gas (GHG) emissions, mostly from the combustion of coal and natural gas to power centralized generators [21]. In addition to greenhouse gas emissions, these power plants are significant sources of other air pollutants, creating short- and long-term health and welfare impacts [24].

Replacement of these centralized conventional generators with renewable energy sources or low-pollution technologies has become a key environmental policy, and many US states have adopted renewable portfolio standards (RPS) to incentivize the development of renewable energy sources [14]. These RPS mandates have been remarkably successful, and have brought California to 20% renewable energy by 2012 and puts California on-path for at least 33% renewable energy procurement by 2030 [15]. This build-out of renewable energy is likely to become particularly important if aggressive GHG reduction goals (e.g. 80% below 1990 emissions levels by 2050) are mandated, which will likely require significant electrification of the vehicle fleet and other energy services [16], [17].

Models of energy deployments under low-carbon scenarios have shown that these changes in the grid are possible under a cap-and-trade system or carbon tax, but that the specific mix of technologies is highly subject to system costs for solar, storage, and demand response resources, all of which are rapidly changing [21–23]. The development of these industries presents significant technical and regulatory challenges, but holds the promise of a low-carbon future.

Technical Challenges

While a few renewable energy sources such as hydropower and geothermal energy can be used as *base load* power, their economic potential has been limited and they have not seen the same rapid growth as wind and solar [23]. While wind and solar are readily scalable, their energy output varies significantly over hourly, daily, and seasonal timeframes. Two terms are used to refer to this: variability and intermittency.

Intermittency describes uncertainty in short-term power output, due to gusts of wind or clouds passing over solar resources. While this uncertainty can be very high at small geographic scales, variance from expected value decreases over larger geographic time scales.

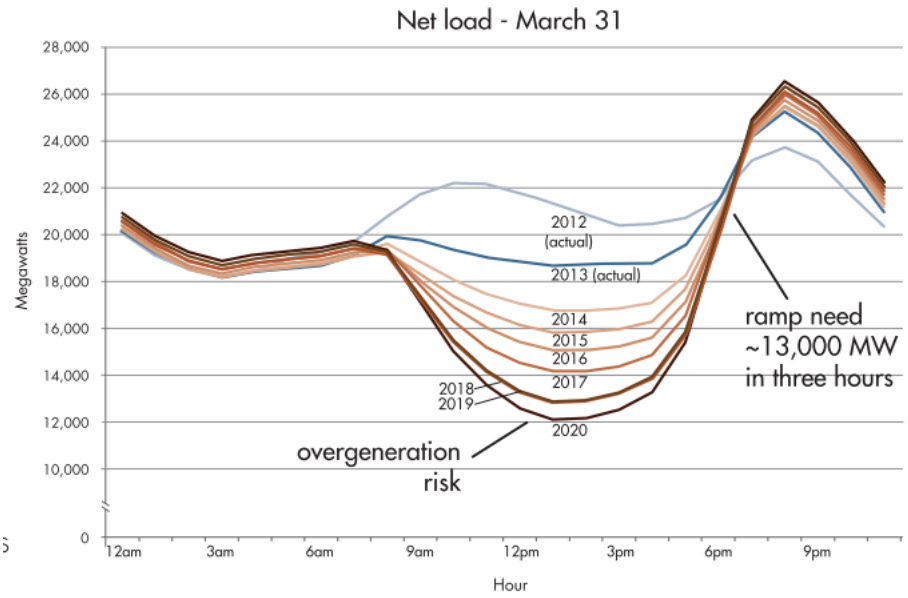


Figure 1.1: Forecasts for a "duck curve" with high renewable penetrations. Source: CAISO [11]

Variability describes the predictable fluctuations in resource availability, e.g. due to the rise of the sun or seasonal patterns in wind availability. Variability on an hourly scale is most significant for solar power, which can have large morning and evening ramps as the sun rises and sets. The mismatch between these diurnal fluctuations and the typical demand curve (which peaks shortly after sundown) can lead to challenges for the system operator, who is faced with a downward morning ramp followed by a long, steep evening ramp when dispatchable resources must be rapidly brought online to maintain grid stability. This problem is summarized in the "duck curve" in Figure 1 which depicts the forecast for net load (load minus must-take generation resources) in future high-renewable scenarios [11, 15].

Overgeneration and Flexibility

The evening ramp along the 'neck' of the duck requires a rapid increase in power generation of roughly 4 GW/hour in the CAISO 2020 forecast [15].This is expected to exceed the current ramping abilities of the generation fleet, resulting in significant overgeneration during the mid-day hours in order to ensure that evening peak load can be served. In a business-as-usual forecast, this would most likely be accomplished by significantly curtailing mid-day solar generation [15, 25], an outcome that would undermine the state's environmental and RPS goals, and result in additional energy costs.

This problem of overgeneration could be introduced by adding system 'flexibility', either in the form of *flexibility down* (increasing system net load, analogous to decreasing generation) or *flexibility up* (decreasing system net load, analogous to increasing generation). The

solutions discussed below (particularly energy storage and load scheduling) achieve these goals through technological means. Additional policy incentives could also incentivize system flexibility response, as CAISO is seeking to do through the introduction of a set of 'Flexible Ramping Products' to provide payments for fast-ramping flexibility [26].

Frequency Regulation

The frequency of the power grid is maintained in a tight band both to ensure the proper operation of electrical devices, and to ensure that the system remains stable. Generator controls are linearized around a local operating region on the power-frequency curve, and significant excursions from this region can result in catastrophic frequency collapse, as the system operating point leaves the linearized region and is unable to adequately adjust power to balance slipping frequency. A summary of the associated controls is provided below; additional detail can be found in [9].

Historically, frequency control has been provided on three time scales: by inertial control, by droop control, and by frequency regulation. Because thermal generators have significant rotational inertia in their turbines, a decrease in frequency will transform rotational kinetic energy into electrical energy, helping balance increased load. Droop controls monitor turbine rotation and provide a proportional/integral increase in fuel consumption to balance slipping frequency. Finally, the system operator can dispatch other generators through automatic generation control to inject additional power into the system (this may be handled through a dedicated frequency regulation market in restructured energy markets).

Renewable power systems, particularly photovoltaics, do not offer the first two of these frequency controls. As they contain no spinning mass but instead synthetically generate an AC waveform from an inverter, renewables increase the frequency sensitivity of the grid to changes in load or generation. Because renewable resources have limited ability to adjust their power output, they cannot provide the same type of droop control as found in conventional generators. This means that the system operator must rely more heavily on real-time dispatch of fast-responding resources to provide frequency regulation, potentially leading to increased energy costs.

This has not been found to be a significant challenge at low renewable penetrations (¡30-40%) [12], but will become a significant challenge at the high RPS levels being called for by California and Hawaii [15]. Addressing this will require tools that can both respond to short-term intermittency, and frequency needs along steep ramps.

Capacity Reserves

System operators seek to ensure that a power grid is able to respond to the unexpected failure of a portion of the generator fleet or transmission network. To do so, they keep a 'capacity reserve' of generators that are either synced with the grid and ready to inject power to the power system on short notice (spinning reserves), or are in standby state and able to rapidly ramp up to provide additional capacity if needed (non-spinning reserves) [9].

The intermittency of renewable resources has led grid operators to require larger capacity reserves to ensure that dispatchable resources are available to meet unexpected changes in supply, e.g. from a cloud bank moving over a centralized solar farm. As renewable penetration has expanded, system operators have increased capacity margins to account for increased intermittency [27], increasing the costs of providing power. Currently, only dispatchable resources such as natural gas turbines are available to provide capacity reserves, but these reserves could also be provided by energy storage or flexible loads [19].

Distributed Generation- Special Concerns

Traditional power generation has been centralized, with the transmission and distribution grids engineered to support centralized control of a small number of large generators producing power for a large number of uncontrolled distributed loads. Distributed generation, particularly rooftop solar generation, has challenged this model, introducing large numbers of uncontrolled distributed generators producing power on the same scale as the distributed loads.

In 2013, the Hawaiian Electric Company (HECO) halted new residential solar net energy metering (NEM) applications on many of its circuits in Oahu, out of concerns of equipment protection in the face of high distributed PV penetration. In the subsequent eighteen months, HECO has moved to a new structure for rooftop solar contracts and revised its equipment requirements for distributed PV installations [28].

When generation on a branch exceeds local consumption, power may flow back through the substation onto the transmission grid. If a distribution link were to fail while power is being fed back to the grid, the isolated circuit must be controlled to minimize the following concerns:

- Transient over-voltage: Excess generation on a disconnected circuit could cause local voltage to rise significantly, known as a transient over-voltage event. Distribution equipment has not typically been designed for this failure mode, and equipment damage could result. To address this concern, HECO is working with NREL to develop a new set of standards for inverter operations [29].

- Unintentional islanding: If generation matches load on the disconnected circuit, the circuit could continue to operate as a self-contained 'island' without control from the rest of the grid. While this could provide beneficial resiliency if planned properly, unintentional islanding may present a risk to line workers or distribution equipment if repairs are needed or the circuit needs to be reconnected with the main [30]. Inverters are currently being designed with anti-islanding controls, but it is unclear whether these operate quickly enough to avoid interfering with automatic reclosers in the distribution grid [31].

- Frequency and Voltage Stability: Intermittency in distributed generation may create frequency and voltage variations that are too localized to be addressed by centralized

generation dispatch systems. For local system stability, it is essential that the distributed generators continue to operate appropriately under local voltage fluctuations, leading to the definition of a standardized 'ride-through' response function which inverters are required to adhere to in order to allow distribution planners to design resilient distribution systems [29].

Designing a cost-effective distributed generation infrastructure presents a unique and challenging exercise in distributed control. While specific standards have emerged to address some symptoms of the problems described above, HECO retains significant limits on its distributed generation deployments [28], and the pathway to Hawaii's proposed 100% RPS is still unclear. Developing a distribution system that is able to provide reliability and security without installing expensive telemetry at each household will require creative design of distributed controls systems, in tandem with electricity rate reform to provide appropriate incentives to emerging technologies and business models.

Planning for Expanded Capacity

The attainment of deep greenhouse gas emissions cuts (e.g. 80% reduction in GHG emissions by 2050) is forecast to be attainable only with significant electrification of our energy infrastructure [16], and concurrent decarbonization of our energy supply in line with the RPS standards discussed here. The model developed by Williams et al in [16] projects a doubling of electricity end-use demand by 2050, requiring significant new investment in capacity for generation, transmission, and distribution. These capacity expansions may have significant environmental impacts in the form of habitat loss or land use change, and accordingly face regulatory or permitting hurdles. These may result in costly delays and environmental impact mitigation measures, as was the case in the Ivanpah solar project [32], or may require collaboration with environmental groups to plot out minimal-impact development plans [33].

Policy Challenges

Many of the solutions described below sit within a specific policy or regulatory framework, which may not be extensible to all power grids. Over the past decades, a number of regulatory structures have emerged for the utility environment, from vertically integrated utilities to fully competitive retail and wholesale energy markets, with utilities serving only to maintain distribution networks [34], [35]. Each wholesale and retail market has unique mechanisms designed to address the integration of renewables, which are continually adjusted as policy goals shift. Detailed discussions of regulatory structure in the electricity space can be found in [34] and [35].

Rate Design

Conventional electricity rates do not present the customer with representative information on the real-time cost of energy, making demand strongly inelastic [35]. Utilities are seeking

to change this by introducing a number of rates that better provide consumers with price signals about the actual price of generating electricity, and thus encourage load shifting to low-cost times. A detailed discussion of rate design is presented in [36], with summaries of key rate structures described below:

- Real-Time Pricing: real-time or dynamic rates provide the consumer with a real-time or nearly real-time (e.g. hourly) rate that changes to match wholesale energy rates. These provide the most information to the consumer about optimum use patterns, but have not been widely deployed due to technological and regulatory hurdles [36].
- Time-of-use Rates: These rates present the consumer with a set of time blocks at differentiated rates, typically 'peak' and 'off-peak' (sometimes also 'part-peak').
- Critical Peak Pricing: Customers pay a low average rate in exchange for being charged a very high 'critical peak' rate during a limited number of events announced shortly in advanced (usually peak cooling load days).
- Demand Charge: Some rate plans, particularly for large commercial customers, include both an energy component ($/kWh) and a demand charge based on peak demand during the billing period ($/kW).

Market Tools

In restructured markets, generation resources are bid into markets for energy production and potentially also ancillary services such as frequency regulation and capacity reserves (described above). As system operators identify weaknesses in the market system, new market mechanisms have been introduced to procure additional services (e.g., energy efficiency, demand response, flexible ramping). The proposed introduction of a set of Flexible Ramping Products into the CAISO market is an example of how market structures are being used to address flexibility needs [26].

1.3 Potential Solutions

The following sections will discuss individual technologies and tools that can be used to address the renewable generation integration challenges listed above. This does not constitute a comprehensive list, and focuses particularly on transmission-scale energy storage, behind-the-meter energy storage, and demand response or load scheduling.

Transmission-Scale Energy Storage

Energy storage systems can offer a variety of services to support the grid, including energy arbitrage, ancillary services, upgrade deferral, increasing reliability, and distribution support services. Both [19] and [37] use engineering estimates to quantify the value of different energy

storage services under current grid conditions, and highlight that the need for storage will grow with the increased penetration of renewable energy.

Storage systems are likely to either be owned by a regulated utility which dispatches them to minimize total cost, or a merchant generator who bids the storage system into energy markets to maximize profits [19]. These dispatch strategies have been modeled using system economic dispatch [38] or by maximizing private storage operator profits when bidding into restructured energy markets [39, 40]. The alignment of public and private optimal dispatch is discussed in [19] and [37], both of which find that storage systems provide more value to the grid than the operators would be currently able to monetize- suggesting that market reforms or subsidies may be needed to support the development of storage.

Current storage capacity is dominated by large pumped-hydro storage systems which are dependent on favorable geography [19], but new storage technology development has focused on modular systems that can be moved into place and sited wherever the grid's needs are greatest [19, 41]. This shift from a small number of large, geographically constrained systems to a large number of smaller storage systems is expected to bring significant opportunities and challenges to the field of power system management [15, 19, 41].

Most existing studies estimate the value of storage based on a specific technology and location, e.g. an existing pumped-hydro storage reservoir as in [39] and [40], limiting their usefulness for guiding future storage deployments. In contrast, [42] identifies the regional arbitrage values of new storage systems by aggregating price nodes for a subset of the PJM Interconnection near New York.

Behind-the-Meter Energy Storage

Some of the largest market potentials for storage systems lie 'behind-the-meter' of the consumer, where storage might be sited next to loads in order to avoid time-of-use rates, demand charges, or critical peak prices [19]. Behind-the-meter systems take advantages of the time-varying rates described in Section 4.3.1 to provide energy savings to the consumer by charging when the grid rate is cheap, and serving the behind-the-meter load when the grid rate is expensive.

Time-of-use rates reflect both the cost of procuring wholesale energy at high costs during high-demand periods, and also of constructing sufficient distribution capacity to ensure peak loads can be adequately served. Behind-the-meter storage systems and the load-shifting services below are uniquely positioned to reduce peak loads, thus allowing utilities to avoid expensive equipment upgrades to serve infrequent peak loads [19], [43].

In the past, behind-the-meter energy storage systems were typically lead-acid battery systems used to provide backup power or off-grid power support [43]. A recent surge of interest in these systems has been spurred by increased deployments of rooftop PV systems and electric vehicles, at the same time that utilities are shifting towards more time-of-use electricity rates and decreasing net-energy metering [36], [44]. These combine to create strong incentives for consumers to store cheap power or zero-cost PV generation to provide for their mobility and household energy needs.

Flexible loads / demand response

Traditionally, the operation of the power market has been invisible to energy loads, and loads only adjusted to infrequent changes in energy rates [35]. Demand response systems have sought to change that relationship, by providing payments to users for reducing their consumption relative to normal levels. These demand response systems have typically have relied on phone calls or messages hours or days ahead of an expected peak load event, asking participants to schedule reduction for a defined period of time (typically 2-6 hours). The transaction costs in this system have typically made it only economical for large commercial loads, and limits its applicability to a few times per year in order to reduce response fatigue.

As more electric loads become connected to the Internet (e.g. thermostats, vehicles, appliances), it becomes possible to schedule or control loads in more applications, at finer time scales. In the context of demand response, this automated switching is referred to as *Automated Demand Response* or ADR, and for our discussion it will mostly be indistinguishable from load scheduling.

Automated demand response or load scheduling can serve to provide both flexibility up (by shifting load away from high-demand periods) and flexibility down (by increasing load in low-demand periods). In this manner, it can be thought of as being very similar to storage systems.

Estimates for a high solar scenario with a 50% RPS in California found that with 5GW of flexible load could decrease the severity of overgeneration incidents by half [15], suggesting that this can be a key tool for addressing renewable energy integration.

Resource Curtailment

The simplest method of handling flexibility needs is arguably the curtailment of excessive generation, in which the system operator or utility turns off generators in order to reduce load. This can create additional costs by incurring startup or shutdown costs, and curtailment is typically compensated with Exceptional Dispatch payments, which are reported to FERC.

We focus on this last example, when the net demand (including exports) is less than the amount that is already committed through must-take renewable resources. In these *overgeneration* scenarios, CAISO requests that committed resources submit additional 'decrement' bids for the cost of having them decrease their generation, and thus alleviate overgeneration through a secondary market. This can compensate generators for costs associated with curtailment, and provide a chance for generators to adjust their bids after the typical market process.

If the decrement market does not result in a sufficient reduction, CAISO calls an *exceptional dispatch* event and takes actions outside of the market, calling generators to request curtailment of generation. These overgeneration events are reported along with other Exceptional Dispatch events to FERC, and the exceptional payments to the generators are compensated through the Load-Serving Entities.

An estimate of the average cost of curtailment on the CAISO grid was developed in [13] based on an analysis of lost revenues in exceptional dispatch events using historical data from 2013-2014, resulting in an estimate of $35/MWh of generation curtailed. This cost is likely to increase if curtailment were to become more common, as renewable energy generators would need to plan for curtailment in their financial analysis and power purchase agreements [15].

Most restructured energy markets feature a set of sequential markets (e.g. day-ahead market, hour-ahead market, real-time market), which act as a set of forwards markets in which participants can update their net position as additional information on load and generation forecasts become available. Additional explanation of sequential markets can be found in [9] and [45].

Power markets differ from financial markets in that there is typically a *gate closure* time, some fixed period prior to dispatch at which point updated bids are no longer accepted, and the system operator works to mitigate market power, clear the market, and schedule generation. Shortening this gate closure period would reduce forecast uncertainty, allowing generators to update their bids to reduce potential overgeneration and reduce the likelihood of curtailment [46].

Adjusting bid limits

To mitigate the profit from exercising market power, most energy markets set limits on the bids accepted from generators. However, when prices clear at this floor, it is not possible to distinguish between the actual costs of shutdown/startup for different generators, creating the potential for uneconomic dispatch. Renewable resources, with zero fuel cost and significant tax credits, may actually have a marginal cost well below $-30/MWh, and so in 2013 FERC approved a change of the CAISO bid floor from $-30/MWh to $-150/MWh, effective April 1 2014 (this will be lowered to $-300 in April 2015) [25]. By reducing the bid floor, CAISO hopes to allow the market to differentiate between renewable resources with a negative marginal cost set by tax credits (totaling up to $115/MWh) or a truly must-run resource that has a much lower marginal cost due to other constraints (equipment repair costs from unscheduled downtime, environmental costs of must-run hydro, etc).

Taken together, we would expect that reducing the price floor will allow prices to go increasingly negative and send a market signal to generators to self-curtail, reducing the need for exceptional dispatch and its associated costs.

Balancing area expansion and transmission build-out

Local pockets of generation imbalance can be ameliorated by strengthening the transmission grid, or by building new transmission lines to connect regions with different load and generation profiles. This may be the least-cost approach to adding flexibility to the grid [12, 46], and is already being tried in some areas.

In 2014, CAISO and PacificCorp created a joint Energy Imbalance Market (EIM) to provide real-time clearing for power transfers between ISOs, partly to mitigate the impact

of increased renewable penetrations [15]. This effectively expands the balancing area to cover a much larger territory, decreasing the variance due to random generation or demand fluctuations. This can be particularly beneficial if the balancing area is large enough to encompass multiple time zones and could thus spread out the peaks of solar generation conventional loads, if sufficient transmission is present [46].

This strategy would not be feasible for geographically constrained grids such as Hawaii's, but where possible it provides an attractive and low-cost solution for well-developed continental grids, and has facilitated high renewable penetrations in European nations which are able to export to neighboring countries [12].

Redesign of conventional generation

If renewable energy targets are modest, the flexibility needs can be met with additional reliance on fast-ramping conventional resources, such as combined-cycle gas turbines [12, 23, 46]. Combined with single-cycle gas turbines, these provide fast-responding and cheap capacity [12] with lower global warming impacts than coal plants. These plants can be further redesigned to be more flexible, allowing for rapid ramping up and down at a slight efficiency penalty [12]. If natural gas prices remain low, this can be expected to provide much of the capacity required to meet a 50% RPS target in California [23].

1.4 Costs

For each of the technical solutions described above, a *levelized cost of flexibility* can be calculated which levelizes capital and operation costs over the lifetime flexibility supplied. These calculations were performed in [12], producing the results presented below. Note that the storage cost projections presented here are higher than those prepared by the Department of Energy in [47], and are 3-4x higher than the DOE's long-term battery storage cost target [43]. The horizontal dashed line at $35/MWh represents the current cost of curtailment derived in [13].

1.5 Concluding Notes

We survey a number of options for increasing the flexibility of the electricity grid in order to be able to meet increased renewable energy integration in line with high Renewable Portfolio Standards. These can broadly be described as enabling the electricity grid to respond to the variability and intermittency inherent in renewable generation systems.

A brief summary of these options, with their corresponding operating time horizons, is shown in Figure 1.3. Each of these present technical, policy, and economic challenges. Integrating them into the power system requires a toolkit which can cost-effectively integrate them to be integrated into energy markets at scale, while respecting the unique physical constraints of both the flexibility resources and of the power network. In the coming chapters,

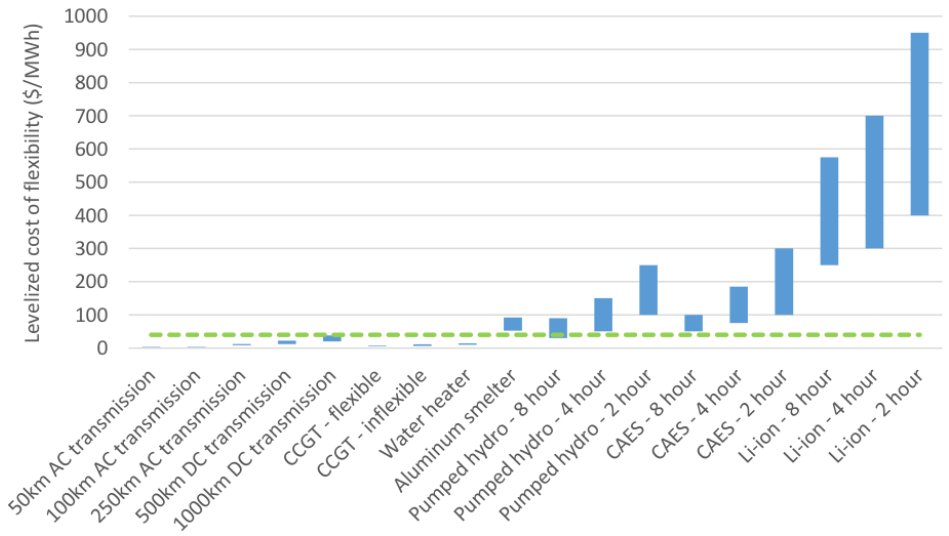


Figure 1.2: Levelized cost of flexibility for transmission, CCGT, demand side solutions, and storage. Costs from [12] and figure from [13]

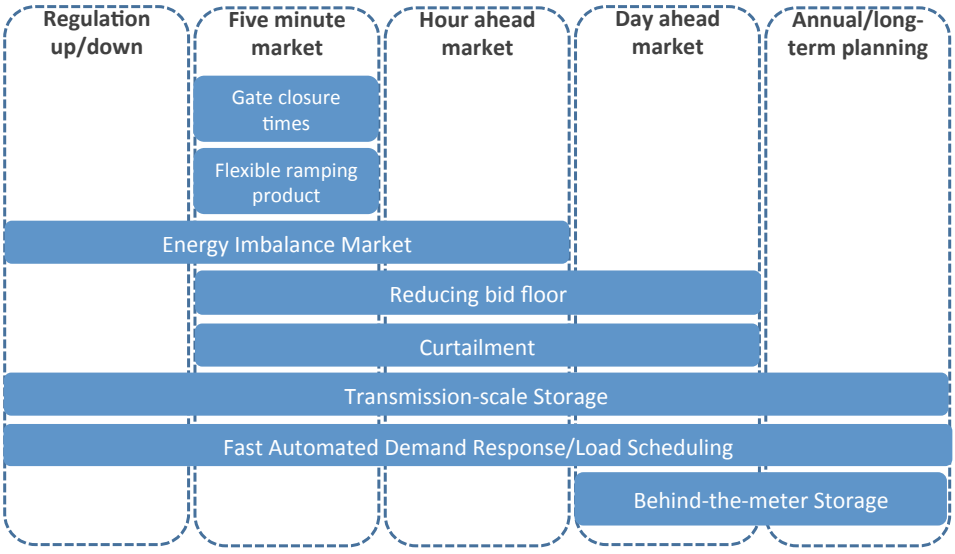


Figure 1.3: Summary of Flexibility Resources

we outline examples of how the combination of convex optimization tools and economic models can be used to show how these emerging resources can be used to integrate these new resources into energy markets.

Chapter 2

Design of Small-Scale Storage Systems

This chapter introduces the reader to the methods used to apply optimization tools to energy systems, by studying a small-scale energy storage system in which a low-power energy source slowly charges banks of batteries and supercapacitors, which are ultimately discharged by a payload with deterministic power draw.

By initially focusing on a small-scale application, we are forced to grapple with system dynamics which are not visible at larger scales. Successful solution techniques for this small-scale problem can be easily scaled up to larger systems such as electric vehicles, stationary storage systems, and mobile devices.

2.1 Introduction

Energy management systems for wireless sensor nodes must deliver reliable power and voltage while being small and low-cost. In prior literature, power reliability is maintained by using active energy management components (voltage regulators, DC/DC converters), but these impose efficiency losses [48] while increasing device cost and size [49, 50]. While the use of an active energy management system allows for a convex formulation of the optimal sizing problem in a two-reservoir battery/capacitor system [51], the passive energy management problem is nonconvex and has not been explored in prior literature.

Building on previous work developing flexible printed batteries [52], capacitors [53], and generators [54, 55] we propose a design which eliminates the need for active power management, and instead passively maintains device performance through optimal sizing of generators, batteries, and supercapacitors. To overcome the nonconvexity in this problem, we employ a Particle Swarm Optimization (PSO) algorithm to identify an optimal design [1], and are able to demonstrate a significant reduction in system size while meeting all payload power requirements.

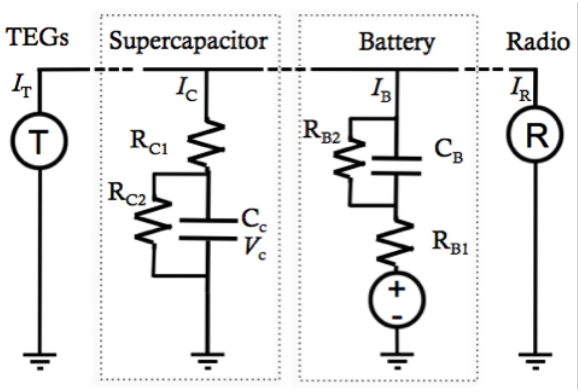


Figure 2.1: Equivalent Circuit Model; the design can be adjusted by changing the number of TEG elements in series in each string, or adding parallel TEG strings, supercapacitors, or batteries in parallel.

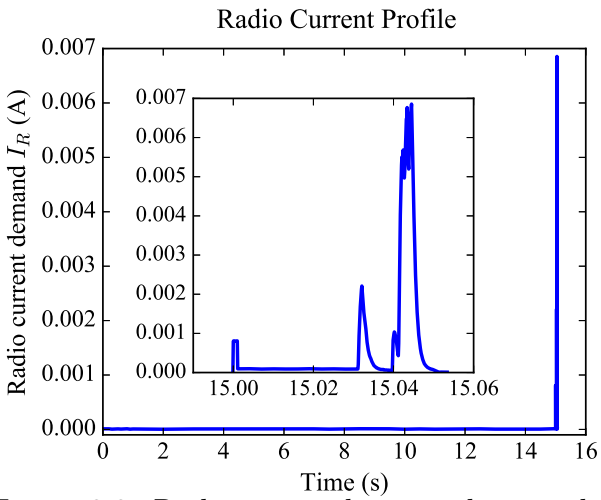


Figure 2.2: Radio power draw, with inset detail of the data acquisition/transmission period.

Novel Contributions

This work extends prior literature on energy management in energy harvesting systems in the following ways:

- Demonstrates a method for designing a passive energy management system without the use of voltage regulators or DC/DC converters, and
- Applies particle swarm optimization to microelectronic energy management system design.

2.2 System Model

We consider a circuit model shown in figure 2.1 in which power is harvested through strings of thermoelectric generator cells (TEGs) in series; the approach we consider can also accommodate photovoltaics or other energy harvesting modes [55]. Charge is stored in parallel banks of supercapacitors and batteries; the components are linked by a main bus at voltage V. A wireless sensor node (WSN) is attached to the bus, and is modeled as a current sink with 15-second cycles with a deterministic current profile. The radio current draw I_R is shown in figure 2.2: a 15-second inactive period is followed by a 50-millisecond data acquisition and transmission load. We seek to design an energy management system which can repeat this load cycle indefinitely.

The system designer chooses the number of TEG elements in series s, number of TEG strings in parallel p, number of batteries b, and number of capacitors c. As the batteries

have a low current capacity and supercapacitors have low energy capacity, we expect that an optimal design may include both storage reservoirs. We assume that each component has a constant per-unit cost, captured in the TEG unit cost α, battery unit cost β, and supercapacitor unit cost γ; the net system cost is thus $\alpha ps + \beta b + \gamma c$.

The system dynamics can be derived through the application of Kirchoff's Current and Voltage Laws (KCL and KVL) and a simple battery model where open circuit voltage OCV is linearized within the acceptable ranges of State Of Charge $\mathrm{SOC_{min}}$ and $\mathrm{SOC_{max}}$. These governing equations and limits on variables are captured below, where italic fonts are used to denote optimization variables:

$$\min \quad \alpha ps + \beta b + \gamma c \tag{2.1}$$

Subject to:

$$\text{KCL for circuit:} \quad pI_T + cI_C + bI_B = \mathrm{I_R} \tag{2.2}$$

$$\text{TEG dynamics:} \quad I_T + \frac{1}{s} V\zeta = \mathrm{T_0 \Delta K} \tag{2.3}$$

$$\text{KVL in Capacitor:} \quad V + I_C \mathrm{R_{C1}} - V_C = 0 \tag{2.4}$$

$$\text{KCL in Capacitor:} \quad \dot{V_c} + \frac{1}{\mathrm{C_c}} I_C + \frac{1}{\mathrm{C_c R_{C2}}} V_C = 0 \tag{2.5}$$

$$\text{KVL in Battery:} \quad V + I_B \mathrm{R_{B1}} - V_B - OCV = 0 \tag{2.6}$$

$$\text{KCL in Battery:} \quad \dot{V_B} + \frac{1}{\mathrm{C_B}} I_B + \frac{1}{\mathrm{C_B R_{B2}}} V_B = 0 \tag{2.7}$$

$$\text{Battery OCV linearization:} \quad OCV - \nu SOC = \mathrm{V_0} \tag{2.8}$$

$$\text{Battery Charging/Discharging:} \quad I_B - I_{BC} - I_{DC} = 0 \tag{2.9}$$

$$\text{Battery State of Charge Conservation:} \quad \dot{SOC} + \frac{\eta}{\mathrm{Q}} I_{BC} + \frac{1}{\eta \mathrm{Q}} I_{DC} = 0 \tag{2.10}$$

$$\text{Inequality Constraints:} \quad b, c, s, p \geq 0 \quad , \quad \mathrm{V_{cutoff}} \leq V \leq \mathrm{V_{max}} \tag{2.11}$$

$$\mathrm{I_{Cmin}} \leq I_C \leq \mathrm{I_{Cmax}}, \quad \mathrm{I_{Bmin}} \leq I_B \leq \mathrm{I_{Bmax}}, \quad \mathrm{I_{Bmin}} \leq I_{BC} \leq 0 \quad , \quad 0 \leq I_{BD} \leq \mathrm{I_{Bmax}} \tag{2.12}$$

2.3 System Design

Conventional Engineering Calculations

We expand on the methods in [49, 50, 56, 57] to provide a baseline estimate of system design using conventional engineering calculations.

To calculate the number of TEG elements in series within each string, and number of TEG strings in parallel, we use the linear TEG dynamics found in equation (3) and consider a constant temperature difference $\Delta \mathrm{K} = 20$. We determine the number of TEGs in series by

dividing the nominal bus voltage by the midpoint voltage of the TEG output curve found in [54]. Similarly, we calculate the necessary number of parallel TEG strings by dividing the average WSN current demand by the midpoint current of a single TEG module.

We calculate the number of supercapacitors required to ensure that maximum payload current $\mathrm{I_{max}}$ can be satisfied without violating $V \geq \mathrm{V_{cutoff}}$. We estimate the minimum number of parallel supercapacitors by dividing $\mathrm{I_{max}}$ by the maximum current a single supercapacitor is able to deliver, $\mathrm{I_{Cmax}}$. We then verify that the combination of coulombic voltage drop and ohmic voltage drop at $\mathrm{I_{max}}$ will not depress the bus voltage below $\mathrm{V_{cutoff}}$.

Because the current limit of the batteries is 2.5% that of the supercapacitors, the battery's contribution to the radio's current demand during peak discharge is assumed to be negligible. Instead, the battery is sized such that it will be able to provide a nominal baseload power, here considered as 10% of $\mathrm{I_{max}}$, and has sufficient capacity to trickle charge the supercapacitor over the cycle should the TEG power drop.

Particle Swarm Optimization

Particle Swarm Optimization is a nonlinear optimization method in which a large number of particles explore the parameter space to identify a low-cost solution [1]. We consider an optimization space in the state variables $SOC(0)$, $V_C(0)$, and $V_B(0)$ in addition to the design variables variables p, s, b, and c. We provide a brief description of our implementation below; additional details can be found in [1] and associated references.

We assign a set of n particles to random locations and velocities in the optimization space, and use the equations above to simulate the system dynamics at each particle's coordinates. If constraints are violated a high penalty is assigned, and the total cost (penalty plus the value of the objective function) is saved for that point. After simulating the system at each particle's location, the location of the lowest-cost point is broadcast to all particles in the swarm. The velocities of all particles are then updated as the weighted sum of the particle's own velocity and the vector to its own best location and the swarm's overall best location. The process is repeated until the swarm converges on a low-cost point, though this may not be the global optimum.

2.4 Results

The results for both design methods are shown in table 2.1 and the resulting bus voltage during the data transmission period is shown in figure 2.3. We assume that a constant per-unit cost for each component, $\alpha = \beta = \gamma = 1$. We see that both designs are feasible, though the nonlinear optimization method converges on a solution which does not use batteries and more closely tracks the system constraints. This contributes to a 55% decrease in the number of components required to power the payload.

A number of other demand profiles were considered, including those in which the TEGs were not able to produce power for a portion of the load cycle. In these scenarios, the

	Engineering Calculations	Nonlinear Optimization
Battery Cells	3	0
Supercapacitors	2	3
TEGs in series	50	22
TEG strings	1	1
Total units	55	25

Table 2.1: System designs resulting from conventional engineering and optimization-based approaches; a 55% decrease in cost is shown.

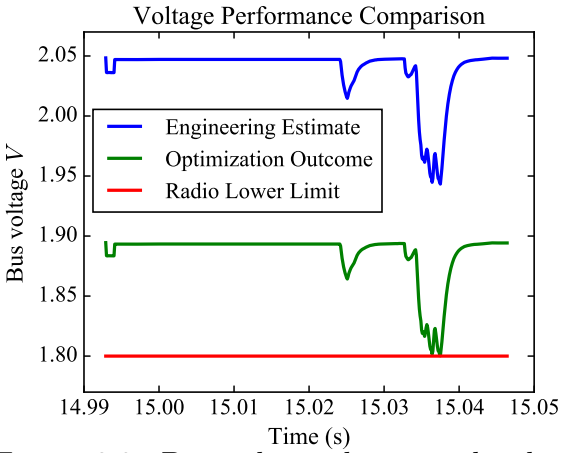


Figure 2.3: Bus voltage during radio dispatch, showing how the nonlinear optimization reduces costs by operating closer to system bounds.

nonlinear optimization resulted in a design in which batteries are added to maintain bus voltage through the drop in generation.

We note that the convergence rate of the nonlinear optimization is dependent on the penalty function, and was found to slow significantly for longer study periods.

2.5 Conclusions

We demonstrate a method for energy management in a wireless sensor node which does not require the use of voltage regulators or DC-DC converters, instead satisfying payload requirements through optimal component sizing. Using particle swarm optimization, we find a design with significant reductions in system cost relative to conventional engineering calculations. While this nonlinear optimization technique does not guarantee a globally optimal design, it can be readily applied to other demand profiles and generation/storage technologies and we expect that similar results may be found in other energy management problems.

2.6 Potential Improvements

The relatively slow convergence of the particle swarm optimization model makes it difficult to scale this optimization approach up to very long time series, large numbers of components, or large numbers of load scenarios. Speeding up convergence would require time-consuming

experimentation and refinement with nonlinear solvers, or finding a convex approximation to the problem.

While the general form of this problem is indeed nonconvex (a bilinear program or nonconvex quadratically constrained quadratic program), the specific problem might still allow for convex relaxations because the problem is highly constrained, and specifically is limited to positive numbers of components.

A variety of techniques were explored in attempting to find a convex relaxation to this problem, including:

- Geometric programming
- Change-of-variables techniques
- Monotonicity Analysis
- Complementarity Programming

Unfortunately, the problem as stated was not compatible with these various tools, particularly because the equality constraint created by Kirchoff's Current Law relaxes into an inequality of the wrong sign, resulting in a nonconvex problem formulation.

The following non-convex approaches offer a promising toolset for this class of problems, but do not offer guarantees of optimality or convergence:

- Sigonomial Programming
- Mixed-integer Programming
- Gradient Descent
- Constraint Satisfaction Solvers
- Nonconvex optimization techniques such as Particle Swarm Optimization, presented here.

We can conclude that both the general class of problems and the specific problem as stated are indeed nonconvex. However, relaxations in which some equality constraints are relaxed into inequalities and combinations of the above tools are applied may offer a relaxation with guarantees of optimality or of convergence.

Chapter 3

Transmission-Scale Storage Sizing and Siting

We next consider a much larger energy system, in which component-level details can be smoothed out through use of power management electronics. By shifting focus, we are able to represent the problem as a linear program, allowing the problem to be solved rapidly and allowing us to evaluate tens of thousands of possible scenarios, over a long time horizon.

This chapter examines in detail the deployment of *transmission-scale energy storage systems*, a promising technology for facilitating renewable energy integration. As discussed in 1, energy storage has been proposed at both the large (transmission scale) and small (behind-the-meter) systems. We here consider transmission-scale storage systems which able to dispatch several enough power to participate directly in energy markets (typically at least 100kW), rather than being controlled by an aggregator. By considering transmission-scale storage systems, we are able to accurately represent the battery as a single unit, and focus on the relationship of the battery size and location with actual electricity prices.

3.1 Introduction

Motivation

The production of electricity from wind and solar energy has become a key element of climate change mitigation strategies, and 29 US states currently have set renewable energy procurement targets through Renewable Portfolio Standards [14]. However, intermittency and variability in these energy sources may lead to curtailment of renewable generation during peak hours and to increased reliance on peaker plants [15] - two outcomes that are environmentally and economically undesirable. Energy storage has been proposed as a technology that can help accommodate the intermittency of the renewable energy systems, while also providing other services such as increasing reliability, deferring upgrade costs, and providing ancillary services [19] [37]. To promote development of new storage systems, the

California Public Utility Commission has mandated that Californian utilities purchase 1.3 GW of storage by 2020 [20], and other regions are considering similar policies.

While many researchers are working on developing low-cost battery chemistries and storage technologies [41], less work has focused on how to design, site, and dispatch this new wave of storage systems.

This study characterizes the optimal sizing and siting of storage over an electricity grid in order to guide policy makers, utility operators, and energy developers. Rather than studying a specific site or technology, we are interested in understanding the impact of location, system efficiency, and reservoir sizing on the profitability of transmission-scale storage.

Prior Literature

The value of energy storage services can be assessed using a variety of methods, including engineering estimates, system economic dispatch models, and simulation of optimal bidding. For vertically integrated utilities, engineering estimates and economic dispatch models can be used to calculate the savings associated with owning storage systems [19,37,38,58]. However, in regions where the energy industry has been restructured to allow open competition, the profitability of a storage system can be calculated using mathematical programming models. These tools can optimize bidding in wholesale energy markets [42, 59], ancillary service (AS) markets [39, 40], and bidding across multiple nodes in a network [60].

Current storage installations are dominated by large pumped-hydro facilities which are limited to favorable geographic sites, but new technology development has focused on modular systems that can be sited wherever the grid's needs are greatest [15,19,41,61]. While this flexibility promises to be an asset to the new storage technologies, the impact of site selection on system profitability has only been examined in [59] and [42], both of which considered a small number of nodes in the PJM market.

The sizing of a storage system's energy reservoir is another important design decision. Unlike conventional generators, the output of a storage system is limited by the capacity of its reservoir (typically described as the number of hours of storage available at peak output, e.g. a 1MW/4h system). While large reservoir capacities are seen in pumped-hydro storage systems due to economies of scale, new storage technologies typically have constant marginal cost, making such large reservoirs much more expensive [47].

Prior literature on sizing storage systems has focused on âĂIJbehind-the-meterâĂİ installations. In these applications, storage is installed to reduce demand charges [19] in a commercial facility, or to smooth intermittency in a renewable energy plant such as a wind farm [62–64], photovoltaic array [65], or concentrated solar power system [66]. When storage is combined with these renewable energy sources, optimal sizing of the storage reservoir allows the plant operator to guarantee contractual energy delivery in the face of uncertain wind or solar forecasts [62, 64, 65].

By contrast, an independent storage system bidding into restructured energy markets would be sized to maximize profits (or minimize utility costs). The impact of reservoir sizing on arbitrage profits is explored through iterative simulation in [59] and [42], but an

optimal size was not identified, and no prior literature has shown a method for endogenously optimizing the reservoir capacity of an independent storage system.

Novel Contributions

This paper builds on prior literature by examining the optimal sizing of a storage system bidding into wholesale energy markets. We examine the arbitrage value for transmission nodes across the grid operated by the California Independent System Operator (CAISO), allowing us to assess the statistical and geographic distribution of storage profits.

This paper addresses gaps in prior literature by introducing the following novel contributions:

1. Simultaneously optimizing reservoir size and system dispatch for energy arbitrage
2. Demonstrating the dependence of optimal reservoir size on storage reservoir cost
3. Presenting storage values for all nodes on the transmission network of an Independent System Operator (ISO)
4. Demonstrating that the storage profits are not uniform or normally distributed across transmission nodes, but rather show a significant tail of high-value nodes
5. Introducing a visualization tool for graphically describing nodal profits

Outline

In Section 2, we outline the assumptions of our model, based on findings in the literature described above. In Section 3 we propose a linear program (LP) that simultaneously optimizes the dispatch and reservoir size for a storage system, and then introduce the data that will be used for our analysis. In Section 4 we present results, first for a sample of nodes in order to validate the model's output, and then for all nodes of the CAISO grid. We examine the impact of system efficiency and reservoir cost on the profitability of the system, relax the assumption that the storage operator acts as a price-taker, and test the statistical distribution of profits across nodes. We conclude by noting limitations to our work.

3.2 Modeling Assumptions

In the formulation proposed below, we examine the sale of energy in the wholesale energy market, and do not consider ancillary services (regulation up/down, capacity reserves, etc.). Other authors have examined the co-optimization of arbitrage and ancillary services [40, 42, 66–68], and the current results could be similarly extended. We limit ourselves to energy arbitrage for several reasons: (i) not all ISOs operate ancillary service markets; (ii) the ancillary service markets are traded at the regional level and would not affect comparisons

between nodes; and (iii) the ancillary services markets have smaller trading volume than the wholesale energy market, and thus prices are more likely to be affected by the addition of storage [19, 59].

We assume that the cost of constructing the storage reservoir can be represented as a constant marginal cost, and that for accounting purposes it can be amortized across the lifetime of the storage system, as \$/kWh/year. This follows the approach outlined in [47], and is representative for the electrochemical battery systems which have seen the bulk of recent development [61]. Economies of scale (decreasing marginal cost of reservoir capacity) are seen in pumped-hydro systems, flow batteries, and underground compressed-air energy storage, and could be integrated into the current formulation as an affine decreasing cost function (resulting in a convex quadratic program).

We report results for a storage system normalized to 1kW power capacity, and consider a variety of reservoir capacities, e.g. 2, 4, or 8 hours of storage. This allows results to be presented on the basis of kWh capacity for comparison with other studies. We assume that our results would scale up to a large-scale system (MW of power and MWh of capacity).

We assume that the storage operator has perfect foresight of energy prices. While this assumption appears generous, it allows us to compute an upper bound on storage profits, which is useful for system planning. We initially assume that the storage system is a price-taker, i.e. that it is too small for its actions to affect the energy price at its trading node. This assumption is then relaxed in Section 3.4 by allowing market prices to respond to the actions of the storage operator. The assumptions of perfect foresight and price-taking behavior have been examined for individual nodes in [60] and [59].

3.3 Formulation

In the following sections we outline the mathematical formulation of our optimization problem, and the data used in our simulations.

Mathematical Formulation

The parameters and variables used in this analysis are defined below, with optimization variables presented in italics. Note that energy flow $c(k)$ into the battery is defined as negative, and energy flow out of the battery to the grid $d(k)$ is defined as positive, consistent with accounting for energy purchases as costs and sales as revenues.

k	Time index, from 0 to time horizon N
Δt	Time step size (hours)
$c(k)$	Energy flow into the battery at time k (kW)
$d(k)$	Energy flow out of the battery at time k (kW)
$\mathrm{P}_{\mathrm{charge}}$	Maximum charge power capacity of the system (kW)
$\mathrm{P}_{\mathrm{discharge}}$	Maximum discharge power capacity of the system (kW)
$\mathrm{c}_{\mathrm{grid}}(k)$	Nodal electricity clearing price ($/kWh)
η_{in}	One-way system efficiency when charging
η_{out}	One-way system efficiency when discharging
$E(k)$	Energy level in reservoir at time k
$\mathrm{E}_{\mathrm{min}}$	Minimum allowable energy level as portion of capacity
$\mathrm{E}_{\mathrm{max}}$	Maximum allowable energy level as portion of capacity
$\mathrm{E}_{\mathrm{init}}$	Starting energy level of the storage system
h	Reservoir capacity (h), in hours of peak discharge
γ	Annualized cost of constructing one kWh of reservoir capacity ($/kWh/yr)

Asymmetry in the power limits and efficiencies can be accounted for by adjusting the appropriate parameters. Using these variables, the optimization problem can be stated as maximizing the net profit from buying and selling energy on the wholesale market, after considering the cost for constructing h hours of storage capacity:

$$\max_{c(k),d(k),E(k),h} \sum_{k=1}^{N} \mathrm{c}_{\mathrm{grid}}(k)\Delta t\,(c(k)+d(k)) - h\gamma \mathrm{P}_{\mathrm{discharge}} \tag{3.1}$$

Subject to the following constraints:

$$\mathrm{P}_{\mathrm{charge}} \le c(k) \le 0 \quad \forall k = 1...N \tag{3.2}$$

$$0 \le d(k) \le \mathrm{P}_{\mathrm{discharge}} \quad \forall k = 1...N \tag{3.3}$$

$$h\mathrm{E}_{\mathrm{min}} \le E(k) \le h\mathrm{E}_{\mathrm{max}} \quad \forall k = 0...N \tag{3.4}$$

$$E(k) = E(k-1) + c(k)\Delta t\eta_{\mathrm{in}} + d(k)\Delta t/\eta_{\mathrm{out}} \quad \forall k = 1...N \tag{3.5}$$

$$E(0) = h\mathrm{E}_{\mathrm{init}} \tag{3.6}$$

$$h \ge 0 \tag{3.7}$$

In this formulation, the signs of the optimization variables $c(k)$ and $d(k)$ are constrained in order to accommodate the inefficiency in the system while preserving linearity in the constraints. Although it is compact, this model advances prior literature by simultaneously calculating both optimal reservoir size h and the optimal storage dispatch pattern $E(k)$.

As the objective and all constraints are affine in the optimization variables $c(k)$, $d(k)$, $E(k)$, and h, this optimization can be solved with standard solvers for linear programs, allowing for rapid simulation of thousands of scenarios.

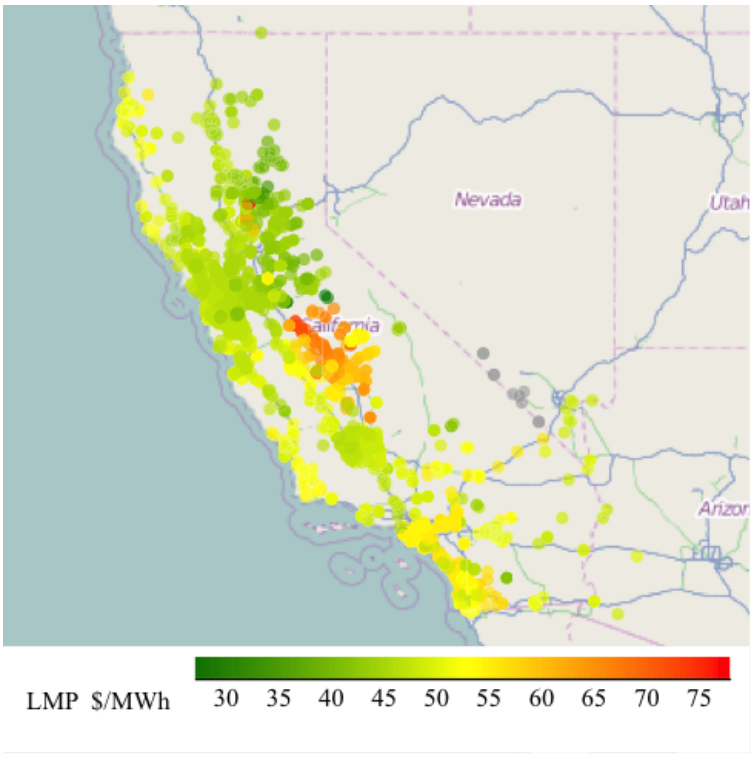


Figure 3.1: Example of Location Marginal Price (LMP) distribution on the CAISO grid. Each circle represents an LMP node on the the CAISO grid. Data from 4PM PDT, August 18 2013

Data

Data on the day-ahead location marginal price (LMP) of energy in the CAISO power grid for calendar year 2013 was collected from the web portal of the of the California Independent System Operator [69].

These LMPs reflect the clearing price at which energy sales are settled on the CAISO grid, and would be the price at which a transmission-connected storage system would buy and sell energy. In the day-ahead market, participants bid in one-hour blocks ($\Delta t = 1$ hour), and all participants bidding at a node receive the hourly clearing price for that node. Data was collected for the 2247 LMP nodes for which location information (latitude and longitude) is available, as shown in Fig. 3.1.

We set $P_{charge} = -1$ and $P_{discharge} = 1$ to represent 1kW of power capacity. For ease of presentation, we will assume that the depth of discharge is not constrained, i.e. $E_{min} = 0$ and $E_{max} = 1$. While many storage technologies have depth-of-discharge limits due to electrochemical or physical constraints, the constraint has the simple effect of adjusting the effective size (and effective cost) of the storage reservoir. For example, a 5-hour reservoir with an 80% depth-of-discharge constraint would show the same optimal trading behavior as

a 4-hour system with no depth-of-discharge constraint (but the 5h system would incur the cost of constructing 1 additional hour of reservoir capacity). Except where otherwise noted, all simulations were conducted with a round-trip efficiency of 90%, which is assumed to be the product of symmetric charging and discharging inefficiencies ($\eta_{\text{in}} = \eta_{\text{out}} \approx 0.95$).

3.4 Results

Because we assume the storage system acts as a price-taker and has perfect price foresight, the results below represent best-case arbitrage profits for a given power-to-energy ratio. For a merchant storage operator to be profitable, the arbitrage profits would need to cover average costs, which we expect will be dominated by the annualized costs of the battery and power system [47]. For this reason, we will refer to the value of the objective function as the âĂIJlong-run profitsâĂİ. We refer to the trading profits excluding the costs of the reservoir construction as the âĂIJshort-run trading profitsâĂİ [6]. Both are presented as \$/kWh/year. For electrochemical batteries that degrade as the system is repeatedly charged and discharged, an important performance metric is profit per cycle (expressed here as \$/kWh/1000 cycles). This is calculated by dividing the short-run trading profits by the number of charge/discharge cycles during the year, and normalizing to 1000 cycles.

Validation

We validate the output of the linear program by evaluating simulation output against the charge/discharge behavior and system size which would be predicted by microeconomic principles, to check that the reported behavior is both economically efficient and profit-maximizing.

As the optimization variables $c(k)$, $d(k)$, and $E(k)$ are linked through the constraints 3.2, we can capture the full system behavior by examining just the reservoir level $E(k)$. In Fig. 3.2 and Fig. 3.3 we plot the charge level for a randomly chosen LMP node from the CAISO dataset (BARRY_6_N001 is shown). Two days in August 2013 are shown, chosen for exhibiting periods of both low and high price volatility.

In Fig. 3.2, the two reservoir sizes are illustrated, obtained by fixing h in the optimization problem and examining the resulting values of $E(k)$. We see that the 6h system charges throughout the morning low-price period, and discharges throughout the evening high-price hours. The smaller system is able to more selectively charge and discharge at extreme price events, but is limited in its ability to capture value from sustained price differences.

In Fig. 3.3 we examine the impact of changing the round-trip efficiency while fixing the reservoir capacity at one hour of storage ($h = 1.0$). At the lowest efficiency (60%), the system only charges once (during the second day), as price differences during the first day are not great enough to outweigh the round-trip efficiency losses. As efficiency increases to 80%, the system is able to take full advantage of diurnal price differences, and begins to arbitrage morning/midday price swings.

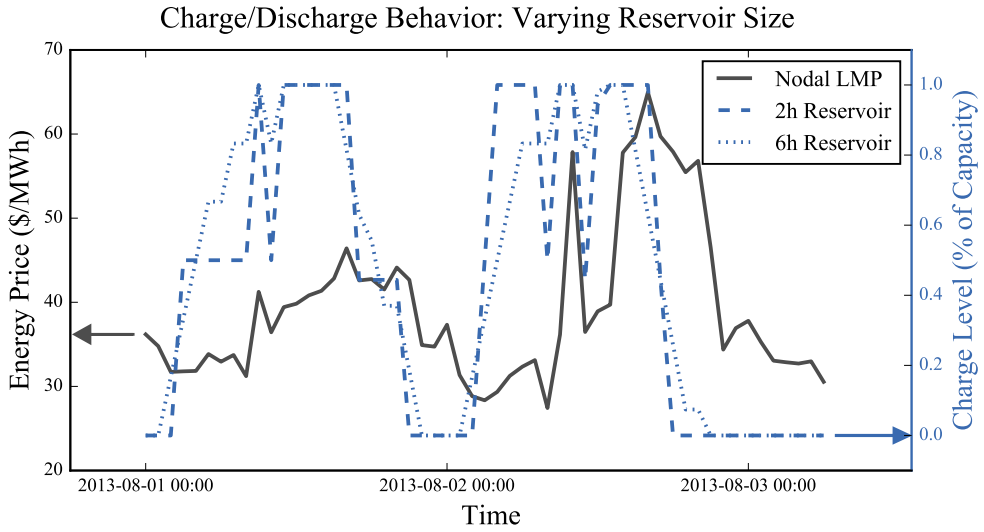


Figure 3.2: Example of the impact of reservoir size h on optimal charging behavior, while system efficiency is held constant. The smaller 2-hour system is able to more quickly charge and discharge, creating larger average price differences for each transaction (and thus greater profits per average cycle). The larger reservoir allows the system to take advantage of sustained price differences- increasing net profits but reducing average profits per cycle. The larger system also incurs greater construction costs, and so optimal system design thus balances reservoir size with construction costs.

This reflects the economically efficient behavior of charging whenever the price difference between two local extrema is greater than the energy loss, $1/\left(\eta_{\text{in}}\eta_{\text{out}}\right)$, i.e. when the short-run trading profits are greater than the transaction cost created by round-trip inefficiencies [6].

In Fig. 3.4 the annualized storage price γ is held constant at \$5/kWh/year while the reservoir size h was constrained to take on a range of values as in Fig. 3.2. The long-term operator profits (including reservoir cost) are plotted against reservoir size h, and are shown to peak at a value of $h = 2.0$, which coincides with the optimal solution of the linear program when h is unconstrained.

At low reservoir sizes in Fig. 3.4, the operator charges and discharges the system at the hours with the most extreme price events (similar to the 2h system in Fig. 3.2). As reservoir capacity increases the operator is able to capture off-peak price differences, but because price differences are smaller the marginal benefit of this additional reservoir capacity decreases. When the reservoir is below optimal size, the increase in trading profits is greater than the annualized costs of constructing the additional reservoir capacity. At optimum, the marginal benefits of adding capacity are precisely equal to the annualized costs of construction the additional capacity, and the operator's profits are maximized [6]. If the reservoir size is increased beyond the optimum, the annualized costs of storage construction overwhelm short-

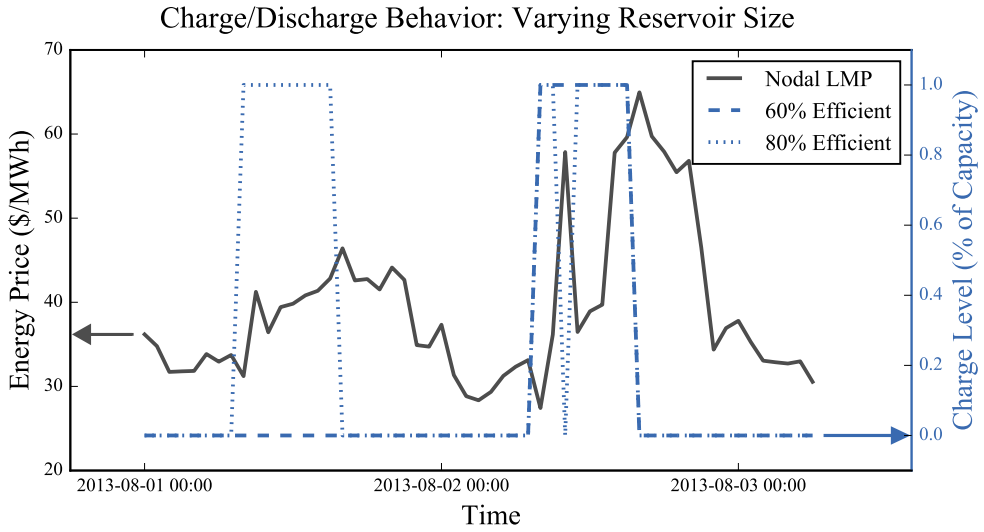


Figure 3.3: Example of the impact of storage system round-trip efficiency on optimal charging behavior, assuming a constant reservoir size. The storage system will only discharge when price differences are greater than the cost of efficiency losses, leading the high-efficiency system to cycle more often.

run trading profits, and the storage system may operate at a net loss.

The shape of this curve will differ for each node, as the returns to increasing storage size are determined by daily, weekly, and seasonal fluctuations in the location marginal price, which will depend on a node's local consumption and congestion patterns.

Sensitivity Analysis: Reservoir Construction Cost

This analysis can be extended by considering a range of reservoir costs, representing the variety of system costs associated with different storage technologies. In Fig. 3.5 the optimal reservoir size is plotted for each node over a range of annualized reservoir costs γ ranging from \$1-\$20/kWh/year, assessed at \$0.10/kWh/year intervals. Note that because the problem is formulated as an LP the resulting curves are not smooth, as the optimum jumps between vertices of the polyhedron defining the feasible region.

As storage price increases we see a monotonic decrease in optimal reservoir size. As shown in Fig. 3.4, the optimal reservoir size occurs when the marginal benefit equals the marginal cost of the additional reservoir capacity: as cost increases this optimum will come at lower reservoir sizes.

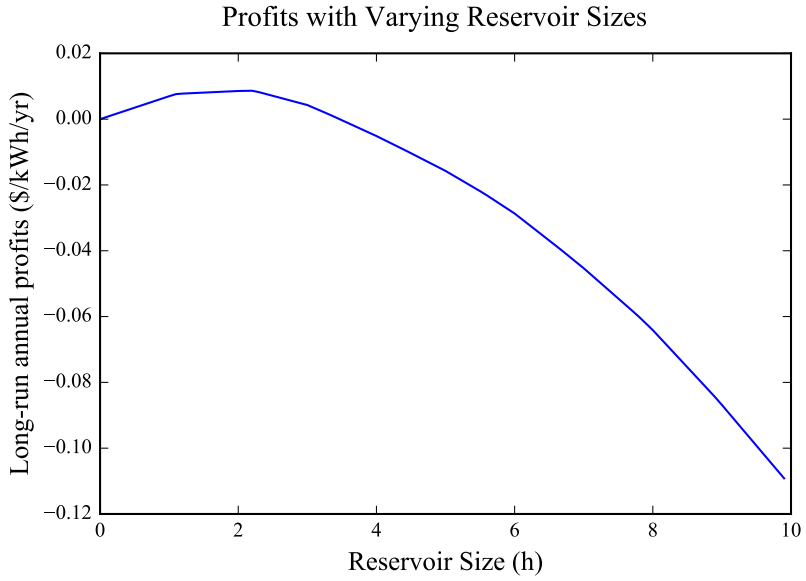


Figure 3.4: Example of the impact of reservoir sizing on long-run operator profits (including the cost of constructing the reservoir). At low reservoir sizes, the operator is not able to take advantage of all arbitrage opportunities. At large reservoir sizes, construction costs outweigh arbitrage benefits. At optimum, the marginal benefit of additional reservoir capacity is balanced by construction costs. Data is shown for the BARRY_6_N001 node in 2013 and assumes an amortized reservoir cost \$5/kWh/year.

Sensitivity Analysis: Storage System Efficiency

A key differentiator between storage technologies is round-trip efficiency, which also has an impact on the optimal dispatch schedule as was shown in Fig. 3.3. To explore the impact of system efficiency on system operation, the optimization was run for all nodes while varying round-trip efficiency from 40% to 100%. This reflects the full range of efficiency values seen in common storage technologies [41, 47].

We examine three metrics: short-run trading profits, number of charge/discharge cycles, and average profits per cycle. The latter two metrics are relevant for many electrochemical battery technologies which have a limited cycle life [47].

For clarity of presentation, in the results below the reservoir size was fixed at $h = 1.0$ (i.e. a 1kW/1kWh system), and results are presented as short-run arbitrage profits excluding reservoir costs. For each node and efficiency, an optimal size could be derived as in Section 3.4.

As efficiency increases, the storage system has fewer losses, increasing profits for a given temporal fluctuation in wholesale energy prices. Simultaneously, the storage operator makes more transactions because they can profitably arbitrage smaller price fluctuations. These

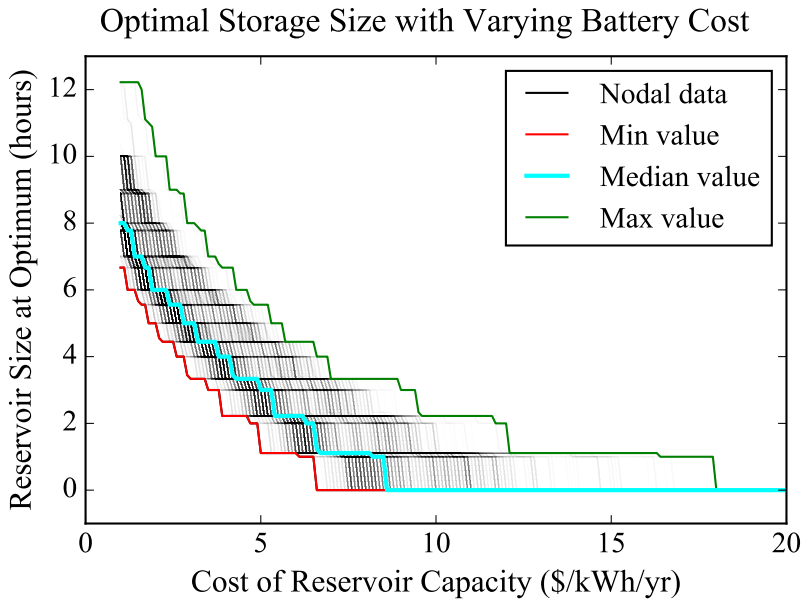


Figure 3.5: Optimal reservoir size in hours of storage capacity with varying annualized construction costs γ, across all CAISO LMP nodes. Optimal size of the storage system decreases when construction costs increase, as the marginal benefits from a larger reservoir are more rapidly offset by construction costs.

effects combine to create a slightly nonlinear increase in profits with increasing efficiency, as shown in Fig. 3.6. A small number of nodes show high profits at low efficiencies due to significant negative price events, when inefficient loads would be paid for their ability to consume more energy.

In Fig. 3.7 the number of charge/discharge cycles is plotted, and shows a slight plateau around 365 charges/year from arbitraging diurnal price differences. There is a rapid increase in cycle count at high efficiencies because the optimization takes advantage of an increasing number of small variations in energy prices. While diurnal and midday/evening price differences occur very regularly, the more frequent trading seen above 95% efficiency is due to minor price fluctuations that may be difficult to predict without perfect foresight.

The per-cycle profits shown in Fig. 3.8 stay fairly constant over the range of 60-90% efficient systems, as the increase in profits is balanced by the increase in number of charge/discharge cycles. However, the per-cycle profits taper dramatically at high efficiencies as the operator chases the small profits associated with frequent hourly price swings.

Comparison with Prior Results

Previous literature estimated the arbitrage value of storage on the CAISO grid to be in the range of \$3-10/kWh/yr [19, 40]. These results are consistent with the median values shown

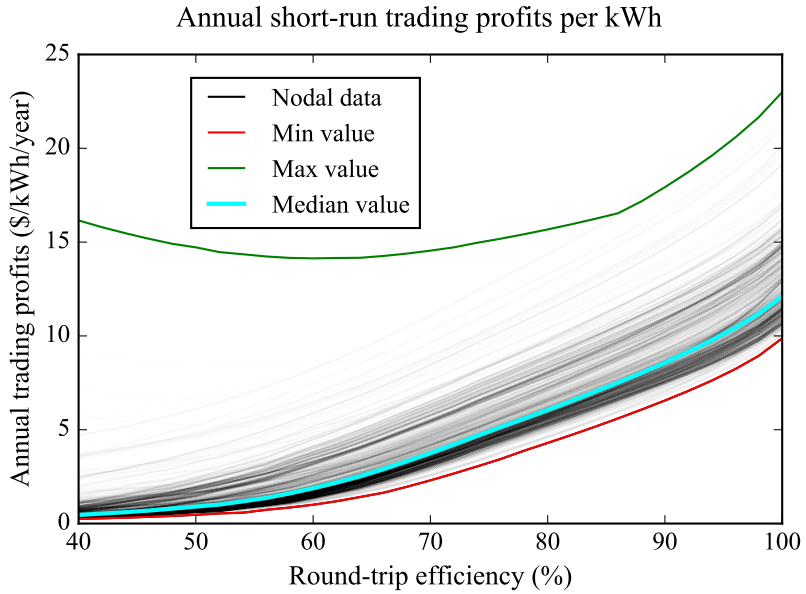


Figure 3.6: Short-run arbitrage profits for each node, plotted at varying efficiencies for a 1kW/1kWh system. Each node is represented by a line plotted at 1% opacity to show the distribution of values at each efficiency.

above, but miss the long tail of high-value nodes. By assessing all LMP nodes, we see that there is a much wider range than previously reported, and that reporting a single value does not sufficiently characterize the distribution of storage value.

Comparing our per-cycle profitability results with the storage system cost and lifetime estimates reported in [47] we find that arbitrage-only profits at high-value nodes could cover the capital costs of pumped-hydro storage, underground compressed-air energy storage, lead-acid batteries with carbon electrodes, and the DOE long-term electrochemical battery cost target. These findings are consistent with existing deployments, which are mostly comprised of pumped-hydro storage [61].

While previous literature has not examined the optimal sizing of a transmission-scale storage system, the optimal sizes found here can be compared with current storage installations. Consistent with our results, high-capacity (8+ hour) reservoirs are only observed in pumped-hydro storage systems, while electrochemical battery systems are most frequently installed in 1-hour or smaller reservoir capacities [61].

Relaxing the Price-Taker Assumption

The model presented above in 3.1 assumes that the LMP values are not impacted by the storage operator's decision, but in reality a storage system used for arbitrage would smooth prices by providing additional supply when prices are high and additional demand when

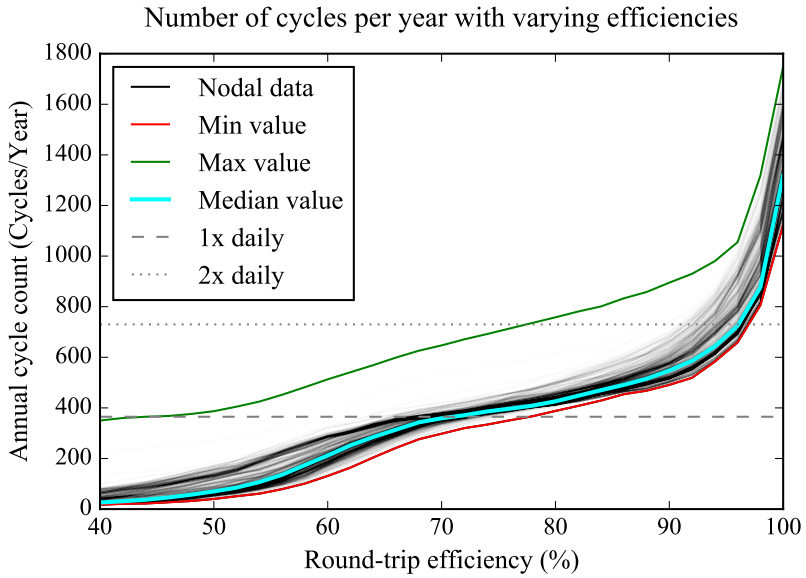


Figure 3.7: Annual cycle count at each node with varying efficiency, 1kW/1kWh system. Thresholds for daily and twice-daily charging are shown, and highlight the significance of diurnal price patterns. Each node is represented by a line plotted at 1% opacity to show the distribution of values at each efficiency.

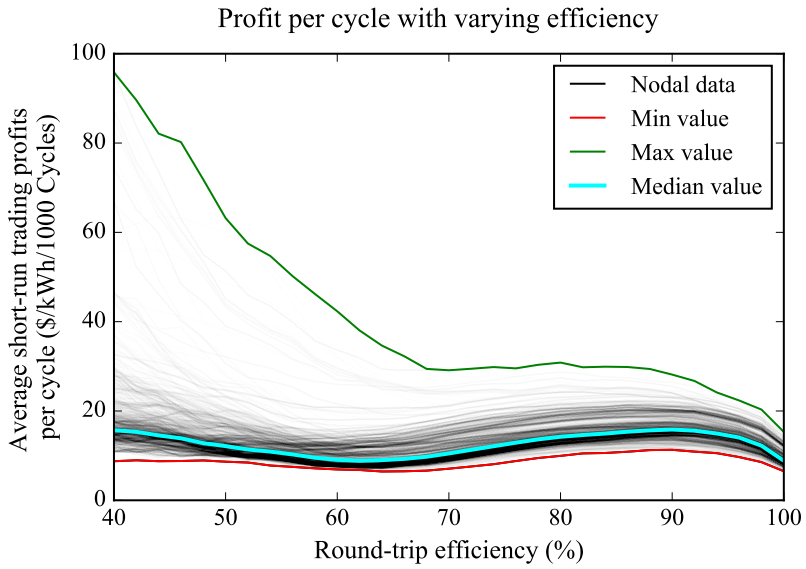


Figure 3.8: Profit per cycle for each node under varying efficiencies, assuming a 1kW/1kWh system. Each node is plotted at 1% opacity to show the distribution of values.

prices are low [6,37]. In the following section we introduce a method for relaxing this price-taker assumption and show that the relaxation reduces profits for each node, but does not affect the relative distribution of profits across the grid.

The impact of market elasticity on arbitrage profits in the PJM market was assessed in [59], using regional trading quantities and clearing prices. However, expanding this approach to nodal LMP calculations is difficult: transaction quantities are not available for individual LMP nodes on the CAISO grid, and local congestion creates nonlinear price elasticity [35].

To overcome these challenges, we propose a simple method to approximate the impact of the storage system on local prices, without assuming that we have a linear elasticity or residual demand curve. We assume that buying energy will drive the market price up by a factor α, and selling energy will decrease prices by the same factor α, regardless of the quantity sold. For example, when $\alpha = 0.10$, we assume that the LMP increases by 10% whenever the storage system is charged, and the LMP decreases by 10% whenever the storage system discharges.

This does not affect the constraints 3.2 and only requires a linear modification to the objective function 3.1, which becomes:

$$\max_{c(k),d(k),E(k),h} \sum_{k=1}^{N} c_{\text{grid}}(k)\Delta t \left((1+\alpha)\, c(k) + (1-\alpha)\, d(k)\right) - h\gamma P_{\text{discharge}} \tag{3.8}$$

For a given value of α, the linearity of the problem is preserved. Because each node faces different transmission constraints and supply curves, no single value of α can be applied for all nodes- at each node, α would need to be estimated from supply disruptions, a study outside of the scope of the present work. Instead, we evaluate a range of values for α in Fig. 3.9, to see the impact of price sensitivity to the price-taker assumption on the profitability of a storage operator under a variety of conditions. As in Section 3.4, we present short-run profits for a 1kW/1kWh system at each LMP node.

As α increases, the arbitrage potential decreases as price variations are smoothed out. However, the relative ordering and distribution of nodes is largely unaffected: our prior conclusions about the distribution and normality of nodes remain unaffected. If price/volume data were available, this assessment could be expanded to include a linear residual demand function at each node, resulting in a convex quadratic program as described in [59].

Distribution of Results

In the results above, storage profits are seen to not be uniform across nodes, but instead exhibit a random distribution. If we *a priori* think of storage profits as the weighted sum of many uniform random variables (congestion, load, renewable energy intermittency, etc.), the Central Limit Theorem suggests that storage values would be distributed normally. However, in the results presented above and in Fig. 3.10, the distribution of storage profits appears to be significantly skewed.

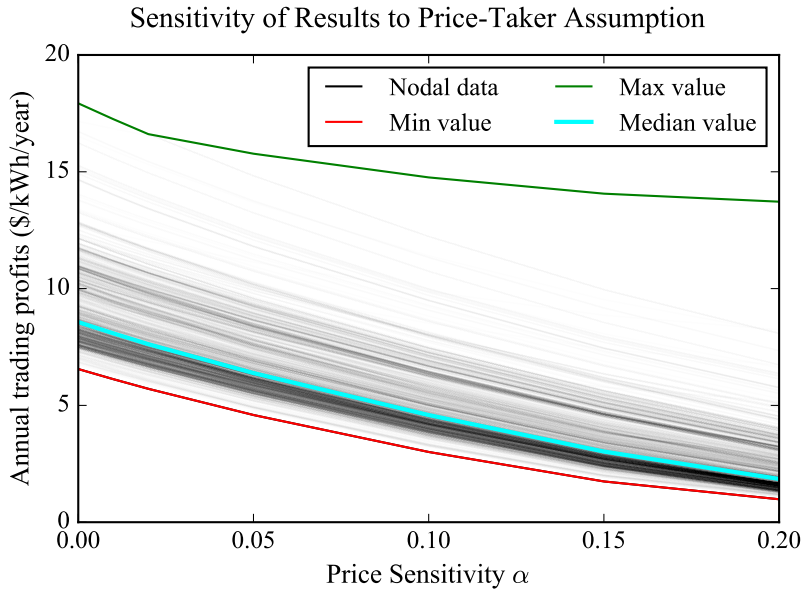


Figure 3.9: Impact of price responsiveness on storage operator profits. We assume that the market price of energy is depressed by a factor α whenever the storage operator sells energy, and is increased by α when the storage system buys energy. While profits decrease, the distribution of the profits across nodes is preserved.

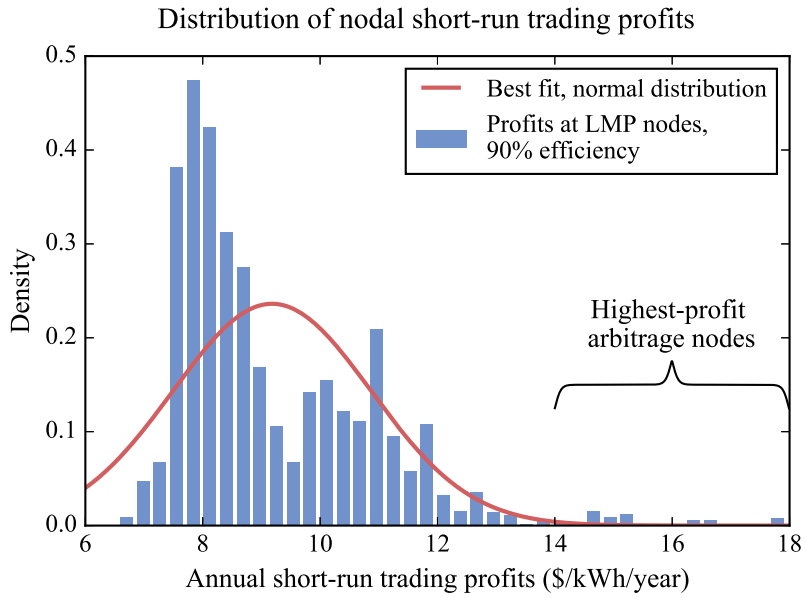


Figure 3.10: Histogram of nodal profits at 90% efficiency for 1kW/1kWh system, with best fit of a normal distribution (estimated with maximum likelihood estimation)

Using both the Kolmogorov-Smirnov and Jarque-Bera normality tests [70] on each of the efficiencies and metrics above, we can reject at the 1% significance level the *a priori* hypothesis of normal distribution. These findings emphasize the significance of high-value nodes in understanding storage feasibility, and how storage site location dictates system profitability.

There appears to be a bimodal distribution in Fig. 3.10 which may indicate additional structure within this distribution; we plan to analyze this in future work.

Geographic Variation in Storage Value

The profitability of storage systems at each LMP was mapped out in a GIS-like web application which the authors built to allow researchers to visualize price data on the CAISO grid [71]. This allows for the geographic localization of high-value and low-value nodes, and assessment of spatial trends. An example of this mapping application is shown in Fig. 3.11, highlighting annual trading profits for the 90% efficiency scenario. Note the clustering of higher-value nodes in the Eureka region and the San Diego foothills, suggesting that congestion in these areas may be partly relieved by the deployment of storage.

3.5 Discussion

Limitations

By construction, this study has only examined transmission-scale storage bidding into the wholesale energy market. Under current regulations, these results do not apply to behind-the-meter storage or vehicle-to-grid services, but could be applied if aggregators were allowed to bid into the wholesale energy market [72].

We only consider bidding into the day-ahead energy market, but the results could be extended to cover sequential markets and ancillary services (AS), as has been done for single nodes in [39, 40, 68]. Since AS markets cover large regions (the CAISO grid has several thousand price nodes but only three ancillary service markets), the inclusion of AS markets would shift the profits of all storage operators in the AS market area without affecting their relative distribution.

The current results reflect a âĂIJbest-caseâĂİ scenario, since they are based on perfect foresight and price-taking behavior using historical market data. The significance of diurnal and morning/evening bidding in most efficiency levels was shown in Fig. 3.7, suggesting that profits are relatively insensitive to relaxing the perfect foresight assumption (as has been done in [59] and [60]). The price-taker assumption was discussed in Section 3.4 and [59] and does not affect our conclusions about optimal sizing, distribution of results, or geographic distribution of high-value nodes.

Only price data for 2013 was analyzed; the inclusion of a larger data set would increase the robustness of the results by introducing variance in natural gas prices, hydropower availabil-

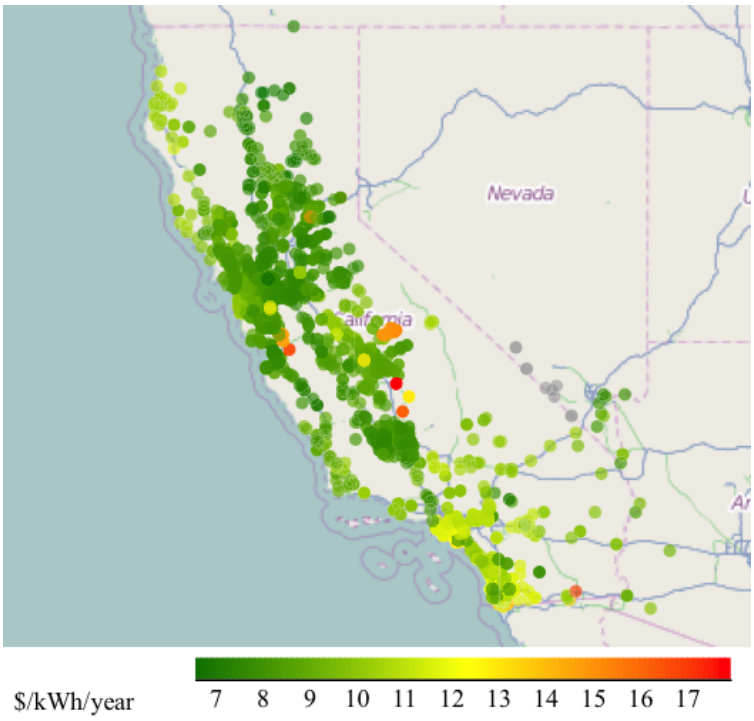


Figure 3.11: Plot of short-run trading profits from energy arbitrage of a 90% efficient 1kW/1kWh system at LMP nodes on the California ISO grid. Broad regions of high-value nodes in the Northern and Southern coastal regions suggest large regions of congestion where storage may have greatest impact.

ity, and renewable energy deployment. The optimal siting of a new storage system would also require forecasting future energy prices under different scenarios of renewable energy penetration.

We have assumed that storage has a constant marginal cost, an assumption which could be relaxed for system which have significant economies of scale (pumped-hydro storage, flow batteries, and compressed-air energy storage) by replacing γ with an affine function of h, which would turn the problem into a convex quadratic program. We have also assumed that the impact of the storage system on energy prices can be represented by a constant factor, a convenient proxy in the absence of historical price/quantity curves. For a specific node, a linear price elasticity or residual demand function could be integrated and result in a convex quadratic objective function. These two modifications modifications (affine economies of scale and linear price elasticity) involve affine terms in separate variables, and thus can both be integrated while preserving the convexity of the formulation.

Conclusion

Using historic price data for 2247 price nodes on the California electricity grid, we utilized a linear program to simultaneously optimize the dispatch and reservoir sizing of a storage system for temporal arbitrage in the day-ahead wholesale energy market. We find that the optimal reservoir size is strongly dependent on installed costs, and that systems with 4 hours of storage or more are only optimal when annualized reservoir costs are below \$10/kWh/year. We explore the dependence of system profitability on efficiency, finding profits of \$7-17/kWh/year or \$10-27/kWh/1000 cycles for a 90% efficient storage system with 1 hour of reservoir capacity. We find a long tail of nodes that are of significantly higher value than have been reported in previous studies, and can reject the hypothesis that storage values are normally distributed. Our revenue estimates show that some existing technologies reported in [47] may be profitable in the highest-value nodes. These results will be of interest to policy makers, utility planners, and storage developers, and will help guide the design and siting of new transmission-scale energy storage systems.

3.6 Potential Improvements

Price Prediction

An extension to this research sought to correlate the value of storage with exogenous covariates (demographics, weather, etc) using a variety of parametric and nonparametric regressors, with the goal of producing a predictor which would guide storage siting for long-term planning or in areas without a wholesale energy market.

These exogenous variables were found to have much poorer predictive quality than a linear regression with the annual average marginal cost of congestion. As the nodal price is composed of the marginal cost of energy (MCE, constant throughout the network), the marginal cost of losses (MCL, usually very small) and the marginal cost of congestion (MCC, often very large), the average MCC over the year ends up being a strong proxy for nodal variation in prices. As the most significant driver of storage revenue is diurnal price swings and MCC is typically very low across the network at night, this average MCC is very closely tied to the value of storage at the node.

This suggests that the proposed line of research has relatively low value for identifying high-value nodes, as simply computing the average MCC forms a good proxy. While the formulation proposed above does provide insight into optimal sizing of the energy system, we were not able to have any ends up being the best proxy for marginal cost of congestion the biggest factor contributing to differences in prices between nodes, the factor most directly related to

Price-taker assumption

The approach used here to modeling the price-taker assumption assumes that all nodes have the same price elasticity. This assumption is likely violated for low-volume nodes, where a small and congested transmission line may result in high-magnitude price swings that would be damped by any flexible resources.

A better model for the impact of dispatch on local prices requires simulating power flow on the transmission network and generating prices endogenously. Unfortunately, there is not a public model of the transmission network, making this difficult to do for real-world systems. An attempt was made to replicate the topology estimation model used by [**?**], but these results were not extended to the full CAISO network.

Instead, the following chapters will consider smaller networks in which the network topology is known. These results are interesting for toy models, but cannot be readily compared against real-world outcomes in the absence of a true network model.

Chapter 4

Strategic Equilibria in Congested Energy Markets

The previous chapter considered the profit-maximizing dispatch schedule for siting individual storage systems. However, as a small number of firms own most of the generation capacity, a more appropriate problem would be of optimizing the dispatch of a *portfolio* of generators at different nodes on a congested network.

Studying this problem requires understanding how generator operation changes power flows and congestion across the network. Rather than assuming that prices are fixed, we will compute prices endogenously by using an optimal power flow model on a simple network with known characteristics.

The previous problem was formulated as a linear program, resulting in solutions which always were at the limits of operation. In this problem, we see how nash-cournot equilibria result in a quadratic program, with interior solutions which are determined by economic criteria rather than physical constraints.

Specifically, this approach uses microeconomic principles as a key component for determining the equilibria which will result from the actions of a set of profit-maximizing actors. This represents a significant improvement over the previous approach, in which prices were fixed and operation was constrained purely by the system's own constraints.

4.1 Introduction

Electricity is unique among commodities, having highly inelastic demand, very limited storage, network flows determined by Kirchoff's laws, and transmission constraints which can isolate consumers from low-cost suppliers. The deregulation of electricity generation in many regions has left the operation of the power grid in the hands of Independent System Operators (ISOs), who are tasked with collecting bids for supply and demand, clearing the market in such a way as to meet transmission and security constraints, and mitigating the use of market power.

However, the characteristics which make electricity unique also make energy markets prone to manipulation, and empirical studies have shown that these markets often operate as oligopolies in which participants affect outcomes by adjusting their bid curves to maximize profits [35], [73]. By modeling strategic equilibria in energy markets, researchers hope to measure social welfare impacts, design regulatory or technical changes which can promote more competitive markets, or identify noncompetitive behavior by comparing models with *ex-post* market outcomes.

A key decision for producers in energy markets is how to respond to uncertainty in supply, demand, and the actions of other producers. Risk-averse producers may hedge their production through long-term contracts, or submit their generation capacity as "must-take" (accepting any price). There have been a number of studies examining the impact of uncertainty on strategic equilibria in energy spot markets, particularly in two-settlement markets, for example [74–76]. However, these approaches can be intractable due to the computational burden required to model uncertainty on a large network using stochastic models [7, 77].

This work advances prior literature by showing how robust optimization can be used to integrate uncertainty into a convex model of strategic equilibrium in a single-settlement Poolco electricity market with network constraints, allowing us to scale our results to large networks while representing the most common form of spot market operation. The results will be of interest to system operators, regulators, and generators in the electricity market, as well as to economists and control theorists studying market design.

Relevant Literature

A number of game theoretic models of strategic competition have been used to model oligopoly behavior in electricity networks, and are reviewed in [7, 77–79]. We highlight three game theoretic models used to identify strategic equilibria in electricity markets: Cournot competition assumes that producers adjust their output to maximize profits [80–82], Bertrand competition assumes producers adjust prices [7], and Stackelberg leader-follower games assume that some firms may have greater decision-making power and serve as market leaders [83, 84]. Of these, Cournot models have gained particular attention for modeling electricity markets due to their mathematical and computational simplicity, as well as their ability to forecast market outcomes [78, 82, 85].

However, since electricity markets are coupled with complex engineered systems these game-theoretic models are not a panacea. Unlike most commodities, electricity markets are built on a transmission network with thousands of nodes [80], have temporal output constraints on generation equipment [86], and are typically structured as a series of sequential markets [87]. To address these issues, a parallel body of literature has sprung up in which engineering models are used to reflect the technical decisions faced by individual producers [88]. These models are often nonlinear, nonconvex, and computationally intractable for modeling the decisions of more than a single producer with a small fleet of generators.

Both game-theoretic and engineering models are challenged by the uncertainty inherent in electricity provisioning: demand is dependent on weather, generation plants may have unplanned outages, and increased penetration of renewable energy sources makes supply uncertain. A variety of techniques from stochastic optimization have been used to model the generators' decision process under these uncertainties [89, 90]: historical energy prices can be used to construct Monte Carlo simulations [91, 92] or to fit parametric probability distribution models to sources of uncertainty [93].

These stochastic approaches can optimize expected profits, but struggle to deal with modeling uncertainty in the hundreds or thousands of nodes which characterize electricity grids. Robust optimization theory [94] and robust game theory [95] provide an alternative approach to integrating uncertainty, by seeking a solution which still performs well under a 'worst case' scenario. While this approach is anticipated to reduce the expected profits for the operators relative to their non-robust actions [96], it can be attractive for risk-averse players as it guarantees profits against uncertainty. These models are also mathematically appealing as they do not require any distributional assumptions on random variables and can preserve the convexity of the optimization problem, allowing the application of efficient solvers which can scale up to handle uncertainty across thousands of nodes.

Robust optimization has previously been applied to game theoretic problems, allowing the modeling of uncertainty in payoff matrices or competitors' strategies [97, 98]. However, robust optimization has only been applied to specific sub-problems in electricity market operation, e.g. the unit commitment problem of the system operator [99–102], nonstrategic bidding as a price-taker [92], strategic equilibrium in a Stackelberg leader-follower game [103], and strategic equilibrium without congestion costs [104].

Novel contributions

We propose a convex formulation, computing the robust strategic equilibrium in an electricity network with congestion and demand uncertainty, and demonstrate it using a sample network.

The following contributions are unique to this work:

- Convex formulation of robust Cournot-Bertrand equilibrium in a single-settlement nodal Poolco electricity market
- Demonstration of the impact of congestion on robust strategic equilibria
- Demonstration of the impact of robust strategies on social welfare outcomes in electricity markets.

To our knowledge this is the first attempt to characterize robust Cournot-Bertrand equilibria in electricity markets on transmission networks, extending the work of [105–107] to incorporate uncertainty and risk-averse producers.

Outline

In Section 4.2 we present the ISO and producer problems as Cournot-Bertrand competition on electricity networks. We formulate the resulting equilibrium as a monotone linear complementarity problem (LCP), which can be solved as a convex QP. We apply the results of [108, 109] to develop a robust LCP and formulate the robust counterpart of the corresponding convex QP. In Section 4.3 we present results for a simple example problem. The impact on producer profits, consumer surplus, and net social benefit is discussed, and we conclude in Section 4.4.

4.2 Problem Formulation

This formulation builds on the work of [105] and [107], obtaining the equilibrium as the solution of a linear complementarity problem, for which a convex robust counterpart is developed. Background on modeling energy markets with complementarity problems is provided in [7] and background on robust optimization theory can be found in [94].

Network Modeling

The power network is modeled by a connected undirected graph $\mathcal{G}(\mathcal{N}, \mathcal{E})$, where $\mathcal{N} := \{1, \ldots, n\}$ is the set of n nodes, and $\mathcal{E} := \{(i,j) : i,j \in \mathcal{N}\}$ is the set of m edges (transmission lines), and $(i,j) \in \mathcal{E}$ means that there is a line connecting buses i and j. Throughout this paper we will assume the standard linear DC power flow model, where lines are lossless, and power flows on the network are governed by a shift-factor (PTDF) matrix $H \in \mathbb{R}^{2m \times n}$ which linearly maps the vector of nodal net injections $r \in \mathbb{R}^n$ to the vector of bidirectional line flows. We denote $T \in \mathbb{R}^{2m}_+$ as the vector of line capacities. Given that nodal injections must sum to zero across the network, we can write the set of feasible power injections as the polytope $\mathcal{R} \subset \mathbb{R}^n$.

$$\mathcal{R} := \left\{ r \in \mathbb{R}^n \mid Hr \leq T,\ \mathbf{1}^\top r = 0 \right\} \tag{4.1}$$

We assume there are $|\mathcal{F}|$ firms, owning generation units at nodes $i \in \mathcal{N}_f \subset \mathcal{N}$, $f \in \mathcal{F}$. Each firm f makes production quantity decisions for its generators $(\{q_i\}_{i \in \mathcal{N}_f})$, where $0 \leq q_i \leq \overline{q}_i$. For simplicity, we assume that there is at most one generation unit per node.[1] Generation costs $C(q_i)$ are assumed to be convex and quadratic.

At each node, we assume a (steeply) decreasing affine inverse demand function $P(x_i)$, where x_i is the quantity demanded at node i. This is a common assumption for relatively inelastic electricity markets [35], and can represent a linearization of a more complicated inverse demand function. It is worth noting that in general x_i is endogenous, and is calculated as $x_i = r_i + q_i$.

[1] This can be achieved in practice by introducing dummy nodes into the network.

The ISO Problem

The ISO controls the import (export) $r_i > 0$ ($r_i < 0$) at each node $i \in \mathcal{N}$ and sets the corresponding locational marginal prices (LMPs). These quantities must satisfy the network feasibility constraints, determined by the set $\mathcal{R}$. The ISO's objective is to maximize social welfare, taken as the aggregated area under the nodal inverse demand functions $P_i(\cdot)$, less the sum of all generation costs $C_i(\cdot)$. Mathematically, the ISO solves the following problem, parametric on the firms' production decisions ($\{q_i\}_{i\in\mathcal{N}}$):

$$\begin{aligned} \underset{r_i}{\text{maximize}} \quad & \sum_{i\in\mathcal{N}} \left(\int_0^{r_i+q_i} P_i(\tau_i)d\tau_i - C_i(q_i) \right) \\ \text{subject to} \quad & 0 = \mathbf{1}^\top r, \ : \ \gamma \\ & 0 \leq T - Hr, \ : \ \mu \end{aligned} \tag{4.2}$$

As in [107], we have excluded the nonnegativity constraints $r_i + q_i \geq 0$, $i \in \mathcal{N}$, by implicitly assuming an interior solution with respect to these constraints. The KKT conditions are as follows

$$\begin{aligned} 0 &= P_i(q_i + r_i) - \gamma - \psi_i, \quad i \in \mathcal{N} \\ 0 &= \psi - H^\top \mu \\ 0 &= \mathbf{1}^\top r \\ 0 &\leq \mu \perp T - Hr \geq 0 \end{aligned} \tag{4.3}$$

The first KKT condition implies that

$$q_i + r_i = (P_i)^{-1}(\gamma + \psi_i), \quad i \in \mathcal{N} \tag{4.4}$$

And consequently,

$$\sum_{i\in\mathcal{N}} q_i = \sum_{i\in\mathcal{N}} (P_i)^{-1}(\gamma + \psi_i) \tag{4.5}$$

This equation represents the aggregate demand function in the network relating the total consumption quantity to the reference node price γ and the nodal price premiums $\{\psi_i\}_{i\in\mathcal{N}}$, which determine the relative value of LMPs. We denote the LMP vector λ as

$$\lambda = \gamma\mathbf{1} + \psi \tag{4.6}$$

To prevent arbitrage between nodes i and j, the corresponding congestion charge must be $\psi_j - \psi_i$.

The Firm's Problem

We assume that generation firms do not anticipate the impact of their production on the congestion prices set by the ISO. We model this 'bounded rationality' as a game where the ISO and generation firms move simultaneously. Similar to [107] we use a mixed Cournot-Bertrand model, where the ISO behaves a la Bertrand, setting locational price differences,

while the generation firms are Cournot players with respect to each other (*i.e.* set quantities), but treat the ISO as a price setter. The reasons for choosing the ISO as a Bertrand player are well discussed in [107].

Each firm chooses its production quantities to maximize profits with respect to the residual demand defined implicitly by (4.5). In this formulation, the reference bus price γ is determined implicitly by the aggregate production decisions of all the generation firms, just as in a regular Cournot game. However, these production decisions and the implied reference node price also depend on the nodal premiums $\{\psi_i\}$ set by the ISO. The resulting problem solved by each firm $f \in \mathcal{F}$ is

$$\begin{aligned} \underset{q_i : i \in \mathcal{N}_f, \gamma}{\text{maximize}} \quad & \sum_{i \in \mathcal{N}_f} (\gamma + \psi_i) q_i - C_i(q_i) \\ \text{subject to} \quad & 0 \le q_i \le \overline{q}_i, \ : \ \nu_i^-, \nu_i^+, \quad i \in \mathcal{N}_f, \\ & \sum_{i \in \mathcal{N}} q_i = \sum_{i \in \mathcal{N}} (P_i)^{-1}(\gamma + \psi_i), \ : \ \beta_f \end{aligned} \tag{4.7}$$

The KKT conditions are as follows

$$\begin{aligned} 0 &= \gamma + \psi_i - \frac{\partial C_i(q_i)}{\partial q_i} + v_i^- - v_i^+ - \beta_f, \quad i \in \mathcal{N}_f \\ 0 &= \sum_{i \in \mathcal{N}_f} q_i + \beta_f \sum_{i \in \mathcal{N}} \frac{\partial (P_i)^{-1}(\gamma + \psi_i)}{\partial \gamma} \\ 0 &= \sum_{i \in \mathcal{N}} (P_i)^{-1}(\gamma + \psi_i) - \sum_{i \in \mathcal{N}} q_i \\ 0 &\le \nu_i^- \perp q_i \ge 0, \quad i \in \mathcal{N}_f \\ 0 &\le \nu_i^+ \perp \overline{q}_i - q_i \ge 0, \quad i \in \mathcal{N}_f \end{aligned} \tag{4.8}$$

We only consider a single market (e.g. spot market) and do not consider optimization across different energy markets (e.g. forward markets or ancillary services), however we will show that it is possible to represent uncertainty with respect to the outcomes of different energy markets.

These assumptions are consistent with other literature [87, 105] and with the approaches used by most ISOs for scheduling hour-ahead and real-time markets, where the computational benefits of the (convex) lossless DC power flow model are important.

Equilibrium Conditions of the Deterministic Game

Aggregating the KKT conditions for the firms' and the ISO's programs yields the equilibrium conditions, which in general form a mixed nonlinear complementarity problem. It becomes a mixed LCP when both the nodal demand functions and the marginal cost functions are linear, as is assumed henceforth.

Let the inverse demand functions and the cost functions be, respectively

$$P_i(x_i) = a_i - b_i x_i, \quad i \in \mathcal{N} \tag{4.9}$$

$$C_i(q_i) = d_i q_i + \frac{1}{2} s_i q_i^2, \quad i \in \mathcal{N} \tag{4.10}$$

where $a_i, b_i, d_i, s_i \geq 0$. We denote $a = \text{vec}(a_i)$, $B = \text{diag}(b_i)$, $d = \text{vec}(d_i)$, $S = \text{diag}(s_i)$.

We denote $L \in \mathbb{R}^{|\mathcal{N}| \times |\mathcal{F}|}$ as the firm-node assignment matrix, where $L_{ij} = 1$ if node i is owned by firm j, and $L_{ij} = 0$ otherwise. We also denote $\beta \in \mathbb{R}^{|\mathcal{F}|}$, where β_i is the dual variable associated with firm i. Also denoting $\sum_{i \in \mathcal{N}} \frac{1}{b_i} = \mathbf{1}^\top B^{-1} \mathbf{1} = c$, the equilibrium conditions are then

$$0 = \gamma \mathbf{1} + H^\top \mu - d - Sq + v^- - v^+ - L\beta \tag{4.11}$$

$$0 = L^\top q - \beta c \tag{4.12}$$

$$0 = \gamma + \frac{\mathbf{1}^\top q}{c} - \frac{\mathbf{1}^\top B^{-1} H^\top \mu}{c} - \frac{\mathbf{1}^\top B^{-1} a}{c} \tag{4.13}$$

$$0 \leq \nu^- \perp q \geq 0 \tag{4.14}$$

$$0 \leq \nu^+ \perp \overline{q} - q \geq 0 \tag{4.15}$$

$$0 = \mathbf{1}^\top r \tag{4.16}$$

$$0 = a - B(q + r) - \gamma \mathbf{1} - H^\top \mu \tag{4.17}$$

$$0 \leq \mu \perp T - Hr \geq 0 \tag{4.18}$$

Here, (4.11)-(4.15) are the aggregated KKT conditions for the firms' problems, and (4.16)-(4.18) are the aggregated KKT conditions for the ISO's problem. Under the assumption of linear demand functions and quadratic convex cost functions, the firms' and the ISO's programs are strictly concave-maximization problems, so (4.11)-(4.18) are also sufficient. Note that (4.13) can be excluded from the preceding market equilibrium conditions because it is implied by (4.16) and (4.17). This set of equations constitutes a mixed linear complementarity program (mLCP).

We wish to turn these set of conditions into a compact LCP. This derivation closely follows that in [107]. We first write out equations (4.16) and (4.17) as follows

$$\begin{bmatrix} a \\ 0 \end{bmatrix} - \begin{bmatrix} B \\ 0 \end{bmatrix} q - \begin{bmatrix} B & \mathbf{1} \\ \mathbf{1}^\top & 0 \end{bmatrix} \begin{bmatrix} r \\ \gamma \end{bmatrix} - \begin{bmatrix} H^\top \\ 0 \end{bmatrix} \mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{4.19}$$

Rearranging and solving for γ and r yields

$$r = Qa - QBq - QH^\top \mu \tag{4.20}$$

$$\gamma = \frac{\mathbf{1}^\top B^{-1}}{c} a - \frac{\mathbf{1}^\top}{c} q - \frac{\mathbf{1}^\top B^{-1}}{c} H^\top \mu \tag{4.21}$$

where, denoting $\mathbf{E} = \mathbf{1}\mathbf{1}^\top$

$$Q = B^{-1} - \frac{B^{-1} \mathbf{E} B^{-1}}{\mathbf{1}^\top B^{-1} \mathbf{1}} \tag{4.22}$$

We note that

$$QB = I - \frac{B^{-1}\mathbf{E}}{c}, \quad BQ = I - \frac{\mathbf{E}B^{-1}}{c} \tag{4.23}$$

We now consider equations (4.11) and (4.12). We have that $\beta = \frac{L^\top q}{c}$, and that $LL^\top = Y$, where $Y_{ij} = 1$ if either $i = j$, or if $i \neq j$ but node i and j are owned by the same firm, $Y_{ij} = 0$ otherwise. Using substitution we rewrite equation (4.11) as

$$\begin{aligned} 0 = \mathbf{1}(\frac{\mathbf{1}^\top B^{-1}}{c}a - \frac{\mathbf{1}^\top}{c}q - \frac{\mathbf{1}^\top B^{-1}}{c}H^\top\mu) + H^\top\mu \\ -d - Sq + \nu^- - \nu^+ - \frac{Yq}{c} \end{aligned} \tag{4.24}$$

Collecting terms, using (4.23), and solving for ν^- we get

$$\nu^- = (BQ - I)a + d + (S + \frac{Y}{c} + \frac{\mathbf{E}}{c})q - BQH^\top\mu + \nu^+ \tag{4.25}$$

We denote $N = (S + \frac{Y}{c} + \frac{\mathbf{E}}{c})$, where $N \in \mathbb{S}_+$ is positive semi-definite, and has the following properties

$$N_{ij} = \begin{cases} \frac{2}{c} + s_i, \text{ if } i = j, \\ \frac{2}{c}, \text{ if } i \neq j, \text{ and the units at nodes } i \text{ and } j \\ \quad \text{belong to the same firm,} \\ \frac{1}{c}, \text{ otherwise} \end{cases} \tag{4.26}$$

We can now write out the following LCP

$$w = \begin{bmatrix} \overline{q} - q \\ \nu^- \\ T - Hr \end{bmatrix}, \qquad z = \begin{bmatrix} \nu^+ \\ q \\ \mu \end{bmatrix},$$

$$t = \begin{bmatrix} \overline{q} \\ (BQ - I)a + d \\ T - HQa \end{bmatrix}, \qquad M = \begin{bmatrix} 0 & -I & 0 \\ I & N & -BQH^\top \\ 0 & HQB & HQH^\top \end{bmatrix}$$

where $w = t + Mz$, $w \geq 0$, $z \geq 0$, $w^\top z = 0$. We notice that M is positive semidefinite but not symmetric. Since it is square we can write M as the sum of a symmetric matrix P and a skew symmetric matrix K, such that $M = P + K$

$$M = \begin{bmatrix} 0 & 0 & 0 \\ 0 & N & 0 \\ 0 & 0 & HQH^\top \end{bmatrix} + \begin{bmatrix} 0 & -I & 0 \\ I & 0 & -BQH^\top \\ 0 & HQB & 0 \end{bmatrix} \tag{4.27}$$

Due to the fact that $z^\top Mz = \frac{1}{2}z^\top(M + M^\top)z = z^\top Pz$, we can solve the LCP by solving the following convex QP

$$\begin{aligned} \underset{z \geq 0}{\text{minimize}} \quad & h(z) = z^\top Pz + t^\top z \\ \text{subject to} \quad & Mz + t \geq 0 \end{aligned} \tag{4.28}$$

with any solution z^* solving the LCP(M, t), iff $h(z) = 0$ [7].

Formulating a Robust Counterpart

We wish to identify strategies for the producers that are robust to uncertainty, an increasingly prevalent feature of modern power systems. Three potential sources of uncertainty for a generator are: the parameters of the inverse demand function, the quantity of zero marginal cost renewable generation in the network, and the volume of forward contracts signed by other generation firms. All of these sources can be represented as aggregate uncertainty in the residual demand curve faced by a producers, however we will see that this formulation can additionally capture more general sources of uncertainty.

We seek a robust equilibrium where producers maximize their profits, robust to demand uncertainty, while assuming that other producers are adopting strategies robust to demand uncertainty. A robust optimization problem takes the form

$$\min_{x \in X} \max_{u \in \mathcal{U}} f(x; u) \tag{4.29}$$

which determines the best possible action $x^\star$ under a worst case realization of uncertainty $u \in \mathcal{U}$. As seen in (4.28), the equilibrium solution of the deterministic problem is an LCP which can be formulated as a convex QP. We obtain a robust equilibrium solution by considering a robust LCP, and formulating the robust counterpart to the equivalent convex QP.

A nominal LCP(M, t) has the form

$$0 \leq z \perp Mz + t \geq 0 \tag{4.30}$$

The function $h(z) = z^\top(Mz + t)$ is known as the residual of LCP(M, t), with $h(z) = 0$ iff z solves LCP(M, t). Applying the results of [108] we define an uncertain LCP(u) as

$$0 \leq z \perp M(u)z + t(u) \geq 0 \tag{4.31}$$

where $M(u), t(u)$ are parametric on the realization of a random variable $u \in \mathcal{U}$. A robust solution to the LCP seeks to find a feasible solution z^* which minimizes the residual function $h(z; u)$ under a worst case uncertainty realization u^*. This takes the form

$$\begin{aligned} \min_{z \geq 0} \max_{u \in \mathcal{U}} \quad & z^\top(M(u)z + t(u)) \\ \text{subject to} \quad & \min_{u \in \mathcal{U}} M(u)_i z + t_i(u) \geq 0, \ \forall i \end{aligned} \tag{4.32}$$

In [109], the authors show that this problem is tractable for affine uncertainty sets of the

form

$$\begin{aligned} t(u) &= t_0 + \sum_{l=1}^{L} u_l t_l \\ M(u) &= M_0 + \sum_{k=1}^{K} u_k M_k,\ M_0 \succeq 0,\ M_k \succeq 0,\ \forall k \\ u &\in \mathcal{U} \subseteq \mathbb{R}^{L+K} \end{aligned} \tag{4.33}$$

where L and K are general scalars defining the affine uncertainty set, and $\mathcal{U}$ can take any of the following forms[2]

$$\begin{aligned} \mathcal{U}_1 &= \{u \ :\ \|u\|_1 \leq 1\},\ \mathcal{U}_2 = \{u \ :\ \|u\|_2 \leq 1\} \\ \mathcal{U}_\infty &= \{u \ :\ \|u\|_\infty \leq 1\} \end{aligned} \tag{4.34}$$

While the formulation is general, for exposition we restrict our attention to uncertainty in $t(u)$, such that $M(u) = M$, $\forall u \in \mathcal{U}$. If $\mathcal{U} = \mathcal{U}_\infty$, then (4.32) takes the form

$$\begin{aligned} \min_{z \geq 0} \quad & z^\top (Mz + t_0) + \sum_{l=1}^{L} \|z^\top t_l\|_1 \\ \text{subject to} \quad & M_i z + t_0 - \sum_{l=1}^{L} \|(t_l)_i\|_1 \geq 0,\ \forall i \end{aligned} \tag{4.35}$$

As stated previously we consider uncertainty in the residual demand function, which we treat as interval uncertainty in the intercept of the inverse demand functions at each node. We consider functions of the form

$$\begin{aligned} P_i(x_i; \zeta_i) &= a_i(\zeta_i) - b_i x_i,\ i \in \mathcal{N} \\ a_i(\zeta_i) &= a_0 + \zeta_i a_{il},\ \|\zeta_i\|_\infty \leq 1 \end{aligned} \tag{4.36}$$

Where $a_{il} \geq 0$ is a common belief among producers regarding the bounds of the uncertainty interval at node i. We implicitly assume that all firms have the same belief regarding the uncertainty in the inverse demand function at each node. This assumption is required for the robust LCP model we have used here. Incorporating different beliefs regarding uncertainty to capture the presence of firms with different risk preferences should be possible, although will require a much closer treatment of the individual robust optimization that each firm faces.

Since a only appears in t in the deterministic LCP, it is straightforward to translate uncertainty in a to uncertainty in t, with the resulting robust LCP having the form of (4.35). This is a convex optimization problem, and thus a robust equilibrium solution exists but may not be unique. The problem can be solved using standard convex optimization solvers.

[2]For the case when $\mathcal{U} = \mathcal{U}_2$, the M_0, M_k are restricted to be symmetric matrices.

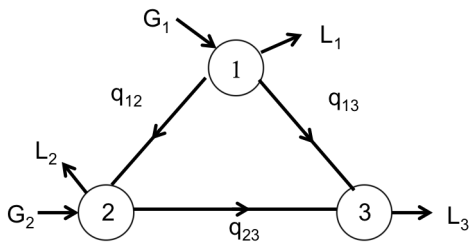


Figure 4.1: 3 Node Network

4.3 Example and Results

We demonstrate our model on an example network, and compare outcomes with conventional (non-robust) Nash-Cournot equilibrium. Networks with up to 300 buses were simulated and were all found to demonstrate qualitatively similar behavior, thus for expositional clarity the simple 3-node network shown in Fig. 4.1 is used here, similar to the network modeled in [110].

Example Network

The three buses $i = 1, 2, 3$ have customers with inverse demand functions $P_i(x_i) = 40 - 0.08q_i$, $i = 1, 2$, and $P_3(x_3) = 35 - 0.05q_3$ \$/MWh (node 3 has greater demand elasticity). Each pair of buses is connected by a single transmission line, and all three lines have equal impedance. The market has two firms $f = 1, 2$ each with a single generator in its fleet; Firm 1's generator is sited at $i = 1$, while Firm 2's generator is at $i = 2$. Both generators have a maximum capacity of $q_i = 1000$MW. Each generator has a constant marginal cost: $d_1 = \$15$/MWh for firm 1, and $d_2 = \$20$/MWh for firm 2. We simulate a congested scenario by imposing a 20MW constraint on the line between nodes 1 and 2, and a 35MW constraint on the line between nodes 1 and 3.

In the examples that follow, we assume that $a_{il} = a_l$, $\forall i$. That is the uncertainty at each node is independent but lies in the same interval. The uncertainty a_l is swept over a range of \$0-15/MWh.

For comparison, the non-robust quantity decisions (i.e. classic Nash-Cournot equilibrium) are also modeled. For each scenario, the non-robust quantity is intercepted with the realized demand curve to produce the price at each node. This simulates the scenario where uncertainty is present, but firms behave as if there is no uncertainty. The system is modeled within Matlab using CVX and the Gurobi solver.

Results and Discussion

The profits for the two generation firms are shown in Fig. 4.2 for the case when the transmission lines are unconstrained and there is no congestion in the network. We plot the results

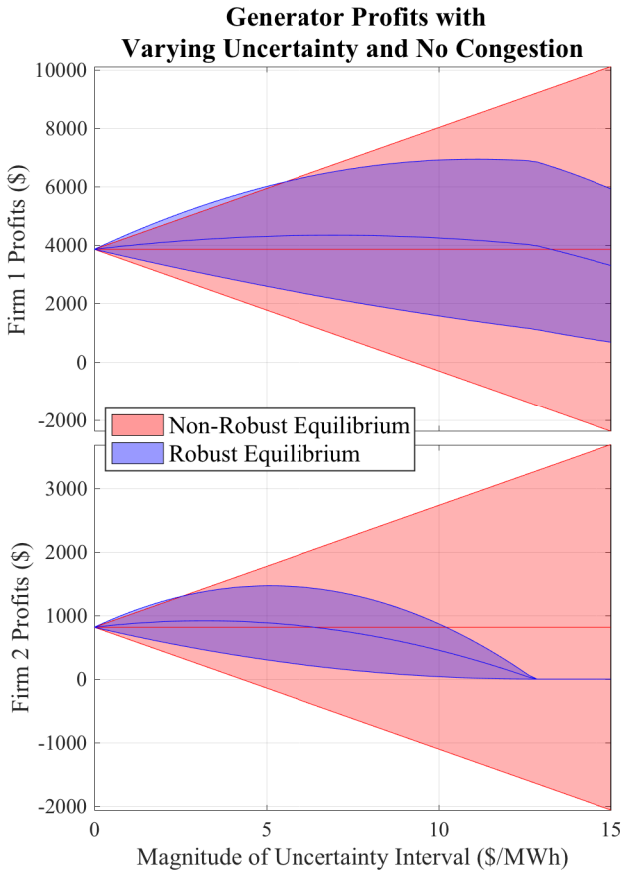


Figure 4.2: Profits of both generation firms under a range of uncertainty intervals. Both robust and non-robust strategies are shown, and the three lines indicate high, expected, and low realizations of demand. The robust equilibrium increases prices for low uncertainty intervals, and limits exposure to downside risk.

for the range of potential uncertainty realizations as a shaded region. These shaded regions are bounded by the profits achieved at the maximum and minimum limits of the uncertainty interval, plotted as thick lines. The results for the nominal value of the uncertainty are plotted as a third solid line through each shaded region. By construction, the uncertainty interval spans all potential realizations of demand and thus all possible market outcomes.

Compared with classic Nash-Cournot equilibrium, we see that the primary goal of the robust optimization is met: the robust strategy always results in higher profits for the worst case realization of demand. For small ranges of uncertainty, we see that the firms actually make *greater* profits at the robust equilibrium than at the non-robust equilibrium, regardless of the level of demand. This can be explained by observing that each firm restricts its output in order to protect itself from low prices, contracting the net supply curve and driving up prices.

Eventually the reduction in demand due to the higher prices offsets the initial gain in

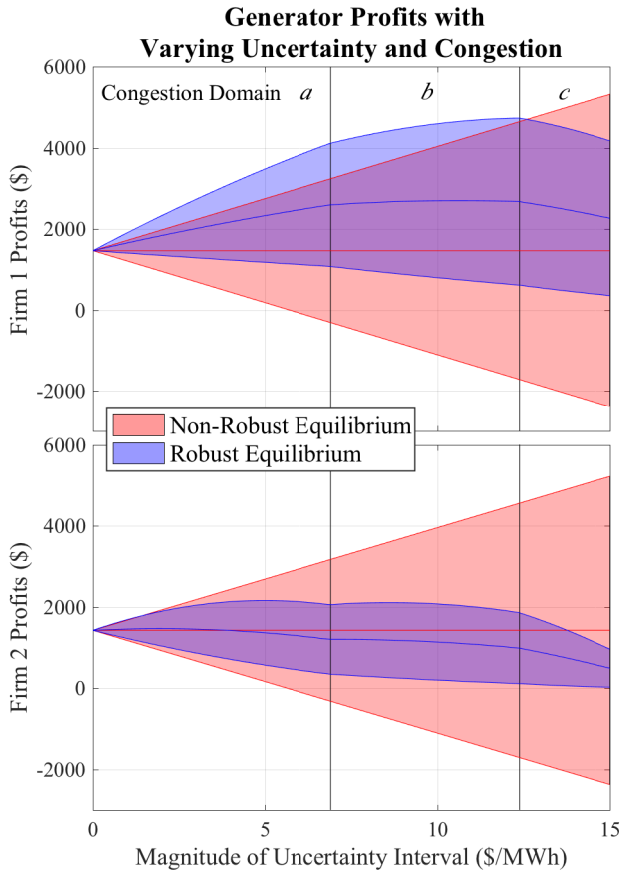


Figure 4.3: Profits of each generating firm when congestion is present on the network. Note that the shape of the curves changes over distinct domains, dictated by the congestion on the network: in domain a line 1-3 is congested; in domain b lines 1-3 and 1-2 are congested, and in domain c only line 1-2 is congested.

profits, and we see that the Nash-Cournot equilibrium results in higher profits in nominal and high demand scenarios. At low demand the robust scenario still guarantees that the generators will not incur a loss, whereas a Nash-Cournot equilibrium can actually result in a net loss for generators as realized prices fall below the marginal cost of generation.

It is important to emphasize that these results assume that all firms follow the same robust optimization behavior, *i.e.* that they have the same belief about uncertainty and the same sensitivity to risk. However, for a less risk-sensitive firm, there is an incentive to increase production in order to increase expected profits. The final equilibrium would be dependent on the firms' risk acceptance, with greater risk aversion driving firms towards the robust optimization, and lower risk sensitivity driving them towards Cournot optimization. This is explored in greater detail in [95].

Network Effects

As firms restrict their production in response to uncertainty, network flows change and can shift the congestion patterns on the network. We can divide the uncertainty range into distinct domains with unique congestion patterns, highlighted in Figure 4.3. Within each domain the residual demand curves for each generator stay constant, and equilibrium follows the principles outlined above for the uncongested case. However as uncertainty increases and the congestion pattern changes, there is a discontinuous shift in the residual demand curve each firm faces, seen as a change of curvature in Figure 4.3.

In our example, comparing Figures 4.2 and 4.3 shows that congestion reduces the profits of Firm 1, but also makes a robust strategy more attractive for low and moderate uncertainty. Firm 2 benefits from congestion rents, but sees profits more threatened by uncertainty in domain a. As uncertainty increases and the line between nodes 1 and 2 becomes congested in domain b, Firm 2 sees greater benefit from uncertainty and the profits of Firm 1 are eroded. This effect is repeated more dramatically as uncertainty increases into domain c. These effects are dependent on the exact network structure, and were found to be particularly complex for larger networks.

Welfare Effects

The impact of the robust strategy on consumers is less nuanced. The consumer surplus is calculated as the area above the market clearing price and below the demand curve, representing the surplus value which consumers would have been willing to pay for electricity [6]:

$$CS = \frac{1}{2}(a - \lambda)^\top x \tag{4.37}$$

The total consumer surplus for the market is shown in Figure 4.4 for competitive, Nash-Cournot, and robust equilibria. As we assume that producers offer a fixed quantity of power into the market, the consumer surplus is invariant to the realized inverse demand function for a given uncertainty interval. When firms restrict their output to be robust to low realizations of demand, prices rise above competitive levels, demand decreases, and consumer surplus drops below the Cournot oligopoly level.

The total efficiency of the market can be measured by its net social benefit: the sum of consumer surplus, producer profits, and merchandising surplus[3] [6]. Building on (4.37), this can be written as

$$NSB = CS + (\lambda^\top q - C(q)) + \mu^\top T \tag{4.38}$$

Since robust behavior restricts supply below Nash-Cournot equilibrium levels, the net social benefit decreases monotonically [6] as shown in Figure 4.5.

[3]This is the rent collected by the system operator in the presence of congestion, and can be shown to be equal to $\mu^\top T$.

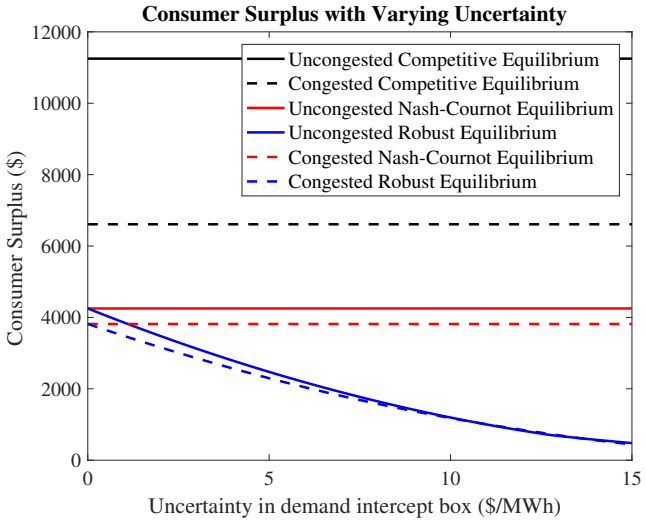


Figure 4.4: Consumer surplus under a number of equilibrium models: perfect competition, Nash-Cournot equilibrium, and Nash-Cournot robust equilibrium. When producers restrict production to be robust to uncertainty, consumers are clearly impacted. Congestion further reduces consumer surplus by introducing congestion charges.

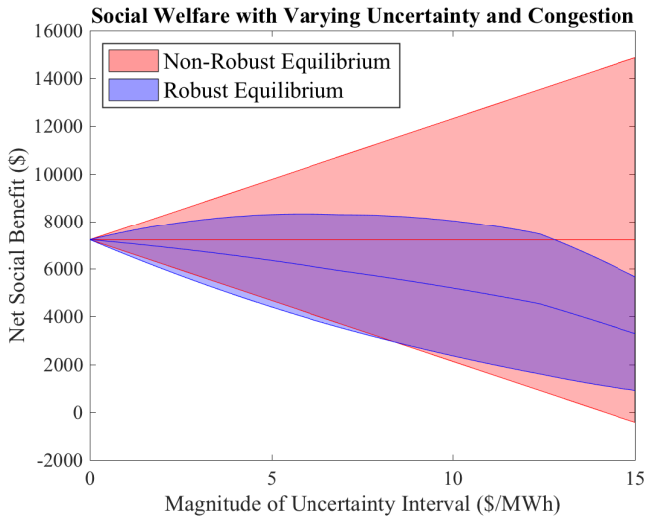


Figure 4.5: Net Social Benefit (sum of consumer surplus, producer profits, and merchandising surplus) under both robust and non-robust equilibrium.

4.4 Conclusion

Electricity markets are particularly susceptible to non-competitive behavior, making it important to understand strategic equilibria in order to inform better market design and policies. The complicated structure of electricity networks, and the many sources of uncertainty in supply and demand, also make it important to have scalable tools for studying the impact of uncertainty on energy markets.

We extend a model of strategic equilibria in electricity networks to include robustness to uncertain demand, reflecting the behavior of risk-averse generation firms. The robust optimization model remains convex, allowing it to be scaled to large power networks. The model is not intended to describe the optimal bidding strategies or bid curves of individual producers, however it provides an efficient way of simulating the impact of uncertainty on market outcomes.

Whereas robust optimization in competitive markets may reduce profits for producers, we see that robustness with small uncertainty intervals *uniformly increases* profits for the generating firms relative to Nash-Cournot equilibrium. The impact of the robust equilibrium on consumers is uniformly negative, as firms restrict their output leading to an increase in prices, similar to that which would be seen under collusive behavior. Thus the "price of robustness" is seen in a reduction of the net social benefit of the market.

Congestion affects different firms unevenly, as it simultaneously creates congestion rents and increases market power for some generators. We show that uncertainty can affect congestion patterns in robust equilibrium, with the exact relationship between generator profits and congestion being dependent on network topology.

These results can be applied to reflect uncertainty in supply due to intermittent renewable electricity generation, by modeling the uncertainty in net load (demand less must-take renewables). The results can also represent uncertainty in the forward contracts signed by other firms, which will contract the residual supply curve in the spot market. By incorporating robustness into the strategic equilibrium, we offer additional insight that producers, utilities, and regulators can use to understand outcomes in real markets.

4.5 Potential Improvements

While this chapter considers the general case of a generator with quadratic costs, this can also be extended to include storage systems, renewable systems, combined solar + storage systems, and aggregated loads. As long as the constraints on these resources can be framed as an affine set, they can easily be integrated into the proposed form without changing the difficulty of the convex optimization problem.

In this chapter, we have still focused on large-scale resources which are dispatched on the transmission network. By constraining ourself to a system with a few hundred or thousand nodes, we can use a centralized solution technique in which a single problem contains all the constraints for the system.

In the next chapter, we will discuss a set of tools which are being used to apply convex optimization techniques to much larger systems in which there may be hundreds of thousands of participants. These *decentalized* optimization techniques move the computation away from a central operator and out to local nodes, which are then brought into consensus on a global optimum. This decentralization can be coordinated through a server which acts as an aggregator, or coordinated by peer-to-peer communication between neighbors in a *fully decentralized* paradigm.

The next chapter will discuss these models in detail, as well as the security risks which they introduce, and how they can be securely executed, using a toy model as an example. The subsequent chapter will then expand on these techniques, applying them to coordinating generation resources on a microgrid with renewable generation, batteries, deferrable loads, and shapeable loads.

Chapter 5

Fully-Decentralized Optimization and Security

The previous chapters have considered optimization models solved on a single server or computer. While simple to express and easy to implement, this centralized model of optimization does not scale well to large numbers of constraints and variables (e.g. scheduling millions of electric vehicles), as most solution algorithms scale super-linearly in the number of variables (typically as $O(n^3)$).

As *distributed* and *fully-decentralized* optimization techniques have been developed for these large problems, smartgrid researchers have proposed moving more of the computation power out to the grid edge- to advanced metering infrastructure (smart meters) and controllable loads (smart devices). While this can allow for significant benefits in computation speed and create robustness to communication dropouts, it turns consumer-level devices into an important part of our cyberphysical infrastructure.

The remainder of this dissertation considers techniques and architectures for securing these decentralized computations in a smartgrid context. As before, we focus on convex optimization techniques, but this time using distributed or fully-decentralized optimization tools. This chapter develops a novel attack detection strategy which can be used in both distributed and fully-decentralized architectures, then describes architectures which can support this. The last chapter will then build on this work to implement a smartgrid control strategy on a blockchain using smart contracts- two novel technologies which take advantage of distributed consensus to secure computations.

5.1 Motivation

With increased computing power, convex optimization tools have gathered attention as a fast method for solving constrained optimization problems. However, some problems are still too large to be solved centrally- e.g. scheduling the energy consumption of millions of electric vehicles.

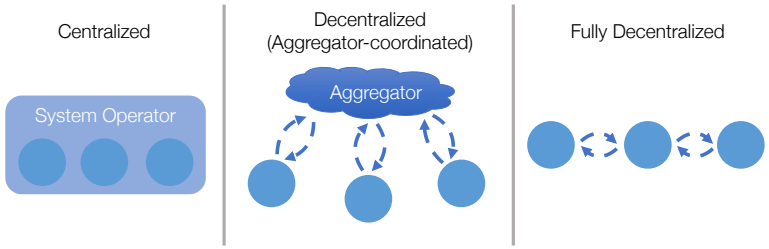


Figure 5.1: Different structures for solving mathematical optimization problems, from left: Centralized optimization, where all details of objective and constraints are held by a central entity. Decentralized optimization (also called aggregator-coordinated optimization) where local nodes hold local objective and constraint information, and an aggregator brings nodes into consensus on shared constraints. Fully-decentralized optimization, where no centralized entity exists but neighbors communicate directly with each other to achieve consensus.

In these applications, *decentralized optimization* techniques break the problem into a set of subproblems which can be rapidly solved on distributed computing resources, with an *aggregator* or *fusion node* bringing the distributed problems into consensus on a global solution. This distribution can yield significant computational benefits, greatly speeding computation times.

A further development on this model, *fully-decentralized* optimization techniques remove the aggregator, and instead let nodes reach consensus with neighbors, in a process which gradually brings the entire system into consensus (see Figure 5.1). This approach significantly reduces communication requirements relative to aggregator-coordinated systems.

The potential for reducing computation time and communication burden makes these distributed and fully-decentralized models particularly compelling in smartgrid applications, where systems may need to scale to millions of devices. However, by moving the computation from enterprise servers to consumer devices, significant new security weaknesses are exposed. This chapter outlines architectures and algorithms for addressing these weaknesses.

5.2 Background Literature

The vulnerability of physical infrastructure to cyberattacks was dramatically highlighted in 2000 when the SCADA system controlling the Maroochy Water Plant in Australia was remotely hijacked, resulting in the release of over one million gallons of untreated sewage [111]. A decade later, the US government developed the *Stuxnet* worm to attack programmable logic controllers which operated uranium enrichment centrifuges managed by the Iranian government [112]; the Iranian government recently launched similar attacsk against targets in Saudi Arabia [113]. A descendant of the Stuxnet worm was used more dramatically by the Russian government to repeatedly cause widespread blackouts in the Ukranian electricity during 2016 and 2017 [114, 115]. Recent incursions into the networks of American electricity

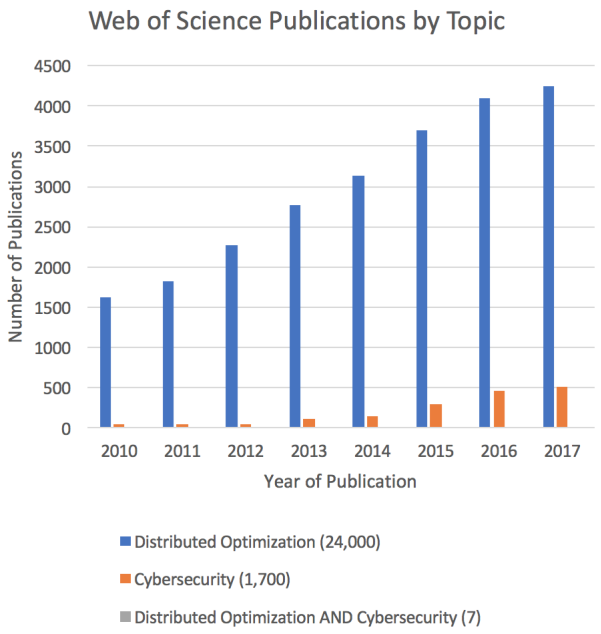


Figure 5.2: Publication frequency for the topics of 'Distributed Optimization', 'Cybersecurity', and the intersection of the two fields, 2010-2017. Legend captions include the total number of publications over this period; the number of publications at the intersection is three orders of magnitude less than each field.

operators by Russian hackers may be laying the foundation for future incursions against U.S. infrastructure [116, 117].

Against this backdrop of cyberattacks used as national security threats, President Obama issued an Executive Order [118] and Presidential Policy Directive [119, 120] directing the executive branch to improve cybersecurity of critical infrastructure. In the face of proven vulnerabilities in consumer-level smart grid devices [121] it seems critical that distributed algorithms for smartgrid controls be designed to be able to identify, localize, and mitigate the effects of these attacks.

Prompted by these concerns, cybersecurity research has increased in recent years, as shown in Figure 5.2. However, this research is far outpaced by the development of new distributed optimization algorithms, and research on the intersection of the two fields - the security of distributed optimization algorithms- is still nascent. Instead, the most cutting-edge research on distributed optimization techniques under adversarial attacks is found not in power systems or optimal control, but in the field of artificial intelligence.

The deployment of distributed machine learning algorithms to very large computation networks or *federated machine learning* platforms (in which consumer feedback e.g. in smartphone apps directly trains an estimator) has led researchers to work on algorithms which can be resilient to software crashes, network dropouts, and adversarial attacks on local nodes.

Most of these machine learning problems are solved with consensus algorithms, in which local nodes solve private problems and reach consensus on a global estimator through iteration with a fusion node (aggregator-based model) or with their neighbors (fully-decentralized model). Although widely used, it has been shown that arbitrary attacks by a single corrupted node can lead consensus algorithms to fail [122,123]- making it critical to understand potential attack vectors, detection algorithms, and mitigation strategies.

As a result, research to combat these approaches has considered a number of methods for developing techniques to allow the consensus algorithm to converge to a best estimate under attack [122]:

- *Filtering techniques* remove outliers from the set of proposed updates [122,124,125], in extreme cases using just the median estimator [126,127] in order to provide tolerance to an attack of half of the nodes. These approaches can be computationally intensive, and may require that the local update functions are all drawn from the same distribution-feasible in statistical estimation, but unlikely in device scheduling.

- *Nonlinear weighting schemes* take advantage of all estimators, but dynamically scale down the impact of suspicious updates. While these are guaranteed to move towards optimum [123,128,129], they have a larger optimality gap than other techniques [130].

- *Round-robin techniques* seek to detect and remove attacked updates by iteratively computing the consensus estimate with different combinations of dropped-out nodes, and using the most stable estimate [125, 131, 132]. This is algorithmically simple but computationally intensive, and is most feasible when the number of attackers is well-bounded.

Additionally, [130] demonstrates the improved convergence of these algorithms when compromised nodes can be switched off.

While many of these algorithms were developed for aggregator-coordinated machine learning problems, many of them can be adapted to fully-decentralized structures, as in the estimation problems studied in [129,133] and the optimization problem studied in [131].

Some of these techniques are beginning to also be deployed in smartgrid optimization problems: [134] demonstrates the hurdle which compromised nodes create for scheduling, much like [135] does for generalized consensus problems.

However, most research on cybersecurity in power system applications has not considered scheduling and optimization problems, but rather state estimation in transmission networks. Liu et al demonstrate in [136] that a single compromised node could introduce an arbitrarily large estimation error, similar to the generalized results in [135]. Subsequent research demonstrated the limitations of attack detection and mitigation strategies under a number of centralized state estimation models [137–139].

These approaches have also been extended to distributed optimization techniques in [132], where a round-robin ADMM method is used to identify compromised nodes; this approach is demonstrated experimentally in [140]. Similar research has also developed fully-decentralized

algorithms for state estimation, with [141] utilizing compressed sensing techniques to recover an estimate of system state when noise injection is bounded, though [142] demonstrates that these approaches still fail when a node is able to inject arbitrary noise.

Novel Contributions

This work advances prior literature by providing an attack detection algorithm for the *alternating direction method of multipliers*, a popular optimization algorithm used for convex optimization and consensus problems. Only requiring that the private objective functions be convex, the attack detection algorithm works for both constrained and unconstrained problems, and both aggregator-coordinated and fully-decentralized architectures. By directly identifying compromised nodes, we bypass many of the objections to the filtering, nonlinear averaging, and round-robin techniques described above, and provide tool which can readily be integrated into computational techniques such as described in [130].

In developing this, this paper provides the following novel contributions to existing literature:

- Outline a taxonomy of attack vectors for decentralized optimization algorithms
- Detail attack detection, localization, and mitigation techinques for the alternating direction method of multipliers
- Demonstrate and verify an algorithm for detecting noise-injection attacks in the alternating direction method of multipliers
- Outline the unique security challenges of fully-decentralized optimization
- Describe potential architectures for security in fully-decentralized optimization

5.3 Outline

In the remaining sections of this work, we outline a taxonomy of methods by which an attacker may seek to compromise a decentralized optimization algorithm, and briefly discuss challenges to detection, localization, and mitigation.

We then develop an algorithm for detecting noise-injection attacks in convex optimization problems solved with ADMM, and present results for a set of simulations with randomly-generated quadratic programs (QPs).

We extend this by considering the unique security challenges presented by fully-decentralized optimization algorithms. We conclude by outlining architectures which can be used to overcome the weaknesses inherent in fully-decentralized architectures.

5.4 Mathematical Background and Notation

Notation

Without loss of generality, we focus on an aggregator-coordinated system with two nodes which compute the x-update and z-update steps in a decentralized optimization problem. Also without loss of generality, we will consider an attack in which the x-update node has been compromised, and use the following notation:

- A, B Constraint matrices in ADMM binding constraint
- c, d Linear cost vectors
- $f(x), g(z)$ generalized objective functions
- H Hessian
- i, j iterates
- k iterate limit
- m, n Number of dimensions of z and x variables respectively
- p Number of binding constraints
- P, Q quadratic cost functions in sample problems
- r, s dimensions in Hessian
- u Scaled dual variable
- w combined variable for centralized solution
- x, z Optimization variables
- y Unscaled dual variable
- $\mathcal{X}, \mathcal{Z}$ The private constraints, knowable only to the compute nodes and not publicly shared
- $x^{\star k}$ The unattacked update, solving $x^k := \text{argmin}_{x \in \mathcal{X}} \ f(x) + \frac{\rho}{2} \|Ax + Bz^{k-1} - c + u_z^{k-1}\|_2^2$
- $\tilde{x}^k$ The attacked signal provided by the z node
- $x^{?k}$ The variable update received by the z-update node, of unknown validity
- $\hat{x}^k$ A best response created by the z-update node which has received a signal perceived as being an attack

Decentralized Optimization

In general, decentralized optimization problems can be cast as:

$$\begin{aligned} x^*, y^* = \text{argmin}_{x,z} \quad & W(x,z)) \\ \text{s.t.} \quad & x, z \in \mathcal{W} \end{aligned}$$

Where x and z further satisfy local problems:

$$\begin{aligned} x^* = \text{argmin}_x \quad & f(x, z^*) \\ & x \in \mathcal{X} \\ z^* = \text{argmin}_x \quad & g(z, x^*) \\ & z \in \mathcal{Z} \end{aligned}$$

We focus in this discussion on iterative methods, where updates $x^k = \text{argmin}_{x\in\mathcal{X}} f(x, z^{k-1})$, $z^k = \text{argmin}_{z\in\mathcal{Z}} g(z, x^{k-1})$ solve local updates in an algorithm which converges to a global solution, i.e. $x^k \to x^*, z^k \to z^*$ as $k \to \infty$.

In addition to computational benefits, this architecture is also advantageous when the local objective functions $f(x)$, $g(z)$ or local constraint sets $x \in \mathcal{X}$, $z \in \mathcal{Z}$ contain *private* information which can not be directly shared with other participants or the aggregator. In this scenario, the updates x^k, z^k do not directly reveal private information, yet still allow the system to converge to the global optimum. This is common in markets and scheduling problems, where participants have private constraints or utility functions which they do not wish to share for economic or security reasons.

Additionally, we assume that some information about the private constraint set is publicly knowable, creating a superset which contains the private constraint set: $\mathcal{X} \subset \mathcal{X}_{\text{pub}}$, $\mathcal{Z} \subset \mathcal{Z}_{\text{pub}}$. Note that as all variable updates are in the private constraint set, they must also be in the public set: $x^k \in \mathcal{X} \in \mathcal{X}_{\text{pub}}$.

The Alternating Direction Method of Multipliers

Although any distributed optimization algorithm may be subject to cyberattack, we specifically consider the *Alternating Direction Method of Multipliers* (or ADMM) algorithm reviewed in [4], which has gained popularity due to its simple formulation and guarantees of convergence for convex problems. The ADMM algorithm is used to decompose a problem with separable objective and constraints:

$$\min_{x,z} \quad f(x) + g(z) \tag{5.1}$$
$$\text{s.t.} \quad Ax + Bz = c \tag{5.2}$$
$$x \in \mathcal{X} \tag{5.3}$$
$$z \in \mathcal{Z} \tag{5.4}$$

Note that only equation 5.2 links x and z.

Adding a penalty term for violations of the linking constraint, we can create an augmented Lagrangean defined in the domain $\{x \in \mathcal{X}, z \in \mathcal{Z}\}$:

$$L_\rho(x, z, y) = f(x) + g(z) + y^T(Ax + Bz - c) + (\rho/2)||Ax + Bz - c||_2^2$$

The standard ADMM algorithm can then be expressed with respect to this augmented Lagrangean $L_\rho(\cdot)$ as:

$$\begin{aligned} x^{k+1} &:= \quad \text{argmin}_{x\in\mathcal{X}}\ L_\rho(x, z^k, y^k) \\ z^{k+1} &:= \quad \text{argmin}_{z\in\mathcal{Z}}\ L_\rho(x^{k+1}, z, y^k) \\ y^{k+1} &:= \quad y^k + \rho(Ax^{k+1} + Bz^{k+1} - c) \end{aligned}$$

or in scaled form:

$$x^{k+1} := \quad \text{argmin}_{x\in\mathcal{X}}\ f(x) + \frac{\rho}{2}\|Ax + Bz^k - c + u^k\|_2^2 \tag{5.5}$$
$$z^{k+1} := \quad \text{argmin}_{z\in\mathcal{Z}}\ g(z) + \frac{\rho}{2}\|Ax^{k+1} + Bz - c + u^k\|_2^2 \tag{5.6}$$
$$u^{k+1} := \quad u^k + Ax^{k+1} + Bz^{k+1} - c \tag{5.7}$$

The iterations stop when:

- The primal residual $r^k = Ax^k + Bz^k - c$ has a magnitude below a threshold ϵ_{pri}, i.e. $||r^k||_2 \leq \epsilon_{\text{pri}}$
- The dual residual $s^k = \rho A^T B(z^k - z^{k-1})$ has a magnitude below a threshold ϵ_{dual}, i.e. $||s^k||_2 \leq \epsilon_{\text{dual}}$

As described in [4] the ADMM iterations will converge to the optimal value of the objective function, and the primal and dual residuals will converge to zero.

In this ADMM formulation, we refer to the "aggregator" (or "fusion node") as the agent responsible for updating u. The aggregator:

1. receives updates from the x-update and z-update steps,
2. computes the u-update
3. broadcasts the updated value of u to the nodes responsible for computing the x- and z-updates.

Consensus problems

ADMM has become a popular tool for solving *consensus problems* in both machine learning and power systems engineering, as both fields deal with problems where computation may be spread across thousands (or millions) of nodes. In a consensus problem, local variables x_i and objective functions $f_i(x_i), i \in \{1, \ldots, N\}$ are united by the constraint that at optimality, the local variables mirror the global variable z:

$$\begin{aligned} \min_{x,z} \quad & \sum_{i=1}^{N} f_i(x_i) d \\ \text{s.t.} \quad & x_i = z, \quad i = 1, \ldots, N \end{aligned}$$

Collected, the ADMM form of this is:

$$\begin{aligned} x^{k+1} := \quad & \text{argmin}_{x_i \in \mathcal{X}_)} \; f_i(x_i) + y_i^{kT}(x_i - z^k) + \frac{\rho}{2}||x_i - z^k||_2^2 \\ z^{k+1} := \quad & \frac{1}{N}\sum_{i=1}^{N} x_i^{k+1} + (1/\rho) y_i^k \\ y^{k+1} := \quad & y_i^k + \rho(x_i^{k+1} - z^{k+1}) \end{aligned}$$

Note that this structure generalizes minibatch stochastic gradient descent, in which the local objectives are the gradients of the local estimators.

As this consensus algorithm is a specific example of ADMM, results for the general ADMM problem will also apply to the ADMM consensus algorithm. For clarity, in this paper we will continue to refer to x- and z-update nodes, even though a consensus problem may have thousands or millions of x-nodes and a single z-update (aggregator) node. Despite this difference in scale, the results we derive for the general 2-node system will still be applicable to the consensus problem.

We will also note that consensus problems are of particular interest in many power system and machine learning problems, as well as in fully-decentralized optimization problems, which include consensus networks.

5.5 Attack Vectors in Decentralized Optimization

When local problems contain private information, the central coordinator can only check $x, z \in \mathcal{W}$ and cannot directly verify that the updates $x^{?k}, z^{?k}$ in fact solve the private optimization problems. In this scenario, a malicious node may submit a distorted update $\tilde{x}^k$ in order to mislead the central coordinator with the goal of creating a sub-optimal solution, an infeasible solution, or prevent convergence of the iterative algorithm.

In this section, we briefly outline potential attack vectors and methods for addressing them. Of these attacks, the most generalizable but difficult to identify is zero-mean noise injection, which we consider in detail below. These attack vectors can be combined, but detection and mitigation efforts will usually address these separately.

Sub-optimal solution

The compromised node solves a modified objective function $\tilde{f}(x)$ resulting in $f(x^k) < \tilde{f}(x^k)$ and consequently $W(\tilde{x}^k, z^k) > W(x^k, z^k)$ as $k \to \infty$. As this does not change the problem structure (e.g. convexity, private constraints), it is difficult to discern from an equivalent problem with slightly modified characteristics, e.g. different consumer preferences. This makes attack prevention a matter of appropriately structuring game-theoretic incentives to avoid malicious distortion. As such, we do not consider this in greater detail here.

Infeasible Private Constraints

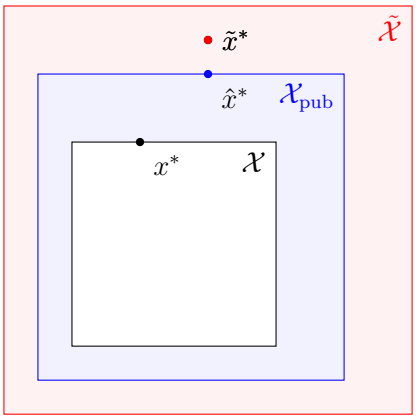
An attacked node may replace the constraints $x \in \mathcal{X}$ with a modified constraint set $x \in \tilde{\mathcal{X}}$ in order to result in an update which is not feasible $\tilde{x}^k \notin \mathcal{X}$, as shown in Figure 5.3. This leads the system to converge to a point which does not reflect actual conditions, e.g. a schedule which is not operationally feasible or a machine learning estimator distorted by false data. Because these constraints may arise from stochastic processes (user preferences, input data) the aggregator cannot directly discern an attack from an unattacked update with different underlying data. As a result, defenses from this attack are limited to cases where the aggregator can bound the support of the problem, i.e. a publicly knowable constraint set $\mathcal{X}_{\text{pub}}$. In this scenario, attacks can be detected when updates lie out of the support region, and the effect of the attack can be mitigated by projecting the update back onto the support region.

Infeasible Linking Constraint

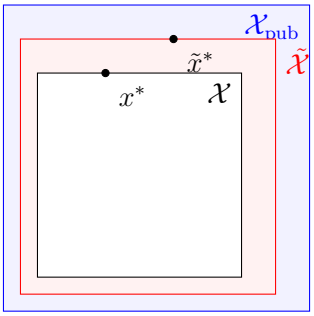
Knowing that a solution to the master problem must satisfy $Ax + Bz = c$, an attacker may present an update which it knows to be unreachable with these constraints, i.e. $\nexists z \in \{\mathcal{Z} | A\tilde{x}^k + Bz = c\}$. In this case, convergence stops as the primal residual $r^k = ||Ax^k + Bz^k - c||_2^2$ remains nonzero. This attack relies on th attacker knowing some bounds on the set reachable by its neighbor $\mathcal{Z} \subset \mathcal{Z}_{\text{pub}}$ in order to shape this attack. Similar to above, detection relies on the same knowledge of the public constraint set, and mitigation relies on projecting the attacked signal back onto the feasible space. If the aggregator has knowledge of the public constraints $\mathcal{Z}_{\text{pub}}$, then detection and mitigation can easily be implemented at the aggregator level for each signal.

Non-Convergence

An attacker may add a noise term which varies with each iteration, i.e. $\tilde{x}^i = x^{\star i} + \delta(i)$. When the noise term is larger than the rate of convergence of the problem, this results in a persistent residual which remains above the convergence tolerance, as shown in Figure 5.4. We examine in detail the case where the noise is unbiased (zero-mean). When the noise is biased, the problem can be considered a combination of a zero-mean noise injection attack and one of the attacks described above, and addressed accordingly.



(a) Attacked system in which detection and (limited) mitigation is possible, as $\mathcal{X} \subset \mathcal{X}_{\text{pub}} \subset \tilde{\mathcal{X}}$.



(b) Attacked system in which detection and mitigation is not possible, as the distorted constraint set is a subset of the publicly known bounds on the constraints, $\mathcal{X} \subset \tilde{\mathcal{X}} \subset \mathcal{X}_{\text{pub}}$.

Figure 5.3: Graphical example of how an attacker may distort the constraint set from $\mathcal{X}$ to $\tilde{\mathcal{X}}$ to create an optimum outside of the truly feasible set, and the limited ability to mitigate these impacts by projecting onto a publicly knowable constraint set $\mathcal{X}_{\text{pub}}$.

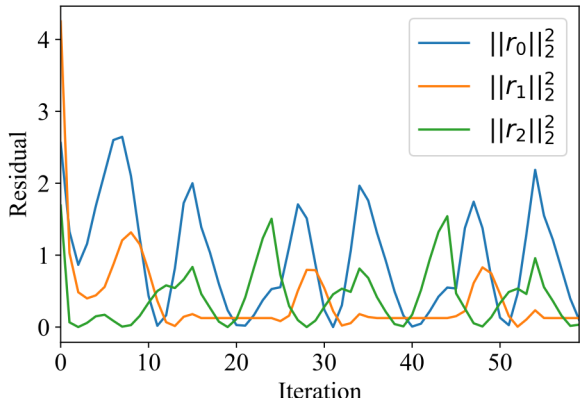


Figure 5.4: Plot of local residuals in a 3-node system under noise-injection attack. Injection of noise prevents problem convergence at nodes across the network.

5.6 Developing a Detection Algorithm for Noise-Injection Attacks

Because the ADMM algorithm relies on private actors computing $x-$ and $z-$updates, these actors can distort the prolem by providing inaccurate updates, with the goal of creating a suboptimal or infeasible solution or preventing convergence. We study the case of distorting the x-update through injection of zero-mean noise, which is intended to prevent convergence of the algorithm. The new update becomes:

$$x^{i+1} := \quad \text{argmin}_{x\in\mathcal{X}} \quad f(x) + \frac{\rho}{2}\|Ax + Bz^i - c + u^i\|_2^2 + \delta(i)$$

Where $\delta(i)$ is a noise term which changes on each iteration. The z-actor or the system aggregator is challenged to identify the attack and take preventative actions before convergence is prevented.

Attack Detection Overview

We play the part of the aggregator or z-actor, and wish to detect an attack of x, using the values of the $x^i, z^i, u^i,\ i = 1 \dots k$ iterates. We rely on *a priori* knowledge that $f(x)$ must be convex, we trust the computations of z^i, and we can verify the u^i values directly by using the other iterates.

We will detect attacks by assessing the convexity of $f(x)$ implied by the x^i iterates, *without being able to directly assess* $f(x)$.

To do this, we will use the $z-$ and u-update values to evaluate the local gradient of $f(x)$, then use these local gradients to construct a finite-differences approximation of the Hessian of $f(x)$. Testing whether this Hessian is symmetric positive semi-definite will then allow us to test the convexity of the x-updates; this is visualized in Figure 5.5. Given our *a priori* knowledge that f(x) must be convex, any updates which result in a nonconvex Hessian must represent an attack.

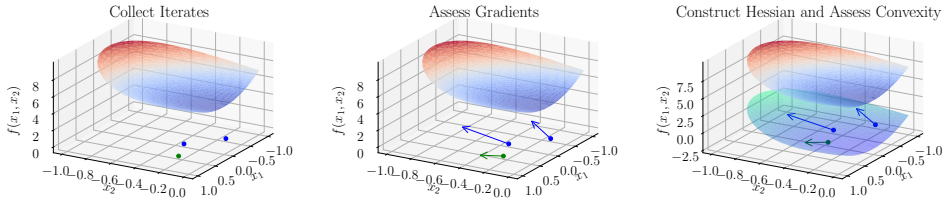
The algorithm progresses in three steps:

- Assess gradient
- Construct Hessian
- Evaluate eigenvalues of Hessian

Assess Gradient

We rely on the proof of ADMM convergence[1] provided by [4], and while we focus on the x-update the same approach can be used for security checks conducted by the x-update node.

[1] specifically building on the *Proof of Inequality A2* on pg 108 of [4]



(a) (b) (c)

Figure 5.5: Conceptualization of the attack detection process for a $n = 2$ toy problem. Although the validator cannot directly assess the private objective function $f(x_1, x_2)$ (shown in red-blue gradient), they are provided with a sequence of iterates x^k, and can to use a subset of these for the detection algorithm, shown here as (x_1, x_2) points with $f(x) = 0$ due to the unknown objective value. Using information from the sequence of iterates, the validator can then assess the implied gradient of $f(x)$ at each of these points. Finally, from these gradients, the validator can estimate the Hessian, and use this to evaluate the convexity of the (unknown) objective function. This can be conceptualized as a local quadratic approximation of the objective function at the reference point, shown in blue-green.

The gradient of $f(x)$ evaluated at x^i is found as:

$$\begin{aligned} 0 &\in \partial L_\rho(x^i, z^{i-1}, u^{i-1}) \\ 0 &\in \partial f(x^i) + A^T + \rho A^T(Ax^i + Bz^{i-1} - c) \\ 0 &\in \partial f(x^i) + A^T(y^i - \rho B(z^i - z^{i-1})) \\ \partial f(x^i) &= -A^T(y^i - \rho B(z^i - z^{i-1})) \\ \partial f(x^i) &= -A^T(\rho u^i - \rho B(z^i - z^{i-1})) \end{aligned}$$

where the last equality simply uses the scaled dual variable.

We will interchangeably use the notation $f(x^i)$ and $f(x)|^{x^i}$ to both indicate the definite evaluation of $f(x)$ at the point x^i. The iterates $i = 1, \ldots, k$ thus give us a sequence of gradient evaluations at points $x^i, i = 1, \ldots, k$. From this sequence of gradients, we wish to construct the Hessian.

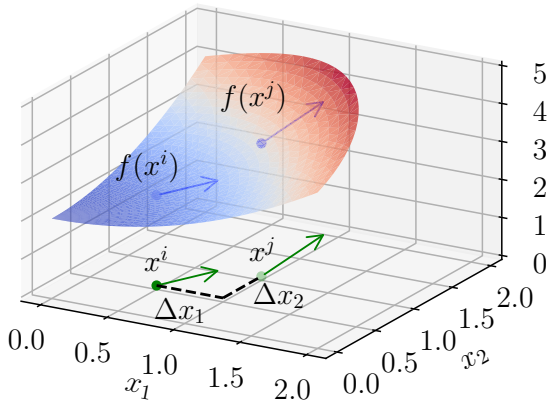


Figure 5.6: Visualization of the relationship between the iterates x^i, x^j, the function surface $f(x)$, and the gradient evaluations $\partial f(x)|^i, \partial f(x)|^j$. While the function surface $f(x)$ and its Hessian cannot be directly evaluated, the change in gradient evaluations between x^i and x^j can be used to derive an approximation of the Hessian.

Construct Hessian

We wish to construct the Hessian matrix

$$H = \begin{bmatrix} \frac{\partial}{\partial^2 x_1} f(x) & \frac{\partial}{\partial x_1 \partial x_2} f(x) & \dots & \frac{\partial}{\partial x_1 \partial x_n} f(x) \\ \frac{\partial}{\partial x_2 \partial x_1} f(x) & \frac{\partial}{\partial^2 x_2} f(x) & \dots & \frac{\partial}{\partial x_2 \partial x_n} f(x) \\ \vdots & \vdots & \ddots & \dots \\ \frac{\partial}{\partial x_n \partial x_1} f(x) & \frac{\partial}{\partial x_n \partial x_2} f(x) & \dots & \frac{\partial}{\partial^2 x_n} f(x) \end{bmatrix}$$

To construct the Hessian, we utilize information gained from the gradient evaluations at each of the iterates to compute a numeric approximation of the local Hessian, using a Taylor series expansion of the Hessian around that point.

Example: 2-Dimension Case

For clarity, we begin with a two-dimensional case, i.e. $x^i = [x_1^i, x_2^i]^T$ and visualized in Figure 5.6. We will consider two points x^i and x^j, and will note the dimensions of these points using the subscripts 1 and 2.

In this case, the change in gradient from point x^i to x^j can be approximated using the following difference equations:

$$\frac{\partial}{\partial x_1} f(x)|^{x^j} = \frac{\partial}{\partial x_1} f(x)|^{x^i} + \frac{\partial}{\partial^2 x_1} f(x)|^{x^i}(x_1^j - x_1^i) + \frac{\partial}{\partial x_1 \partial x_2} f(x)|^{x^i}(x_2^j - x_2^i)$$
$$\frac{\partial}{\partial x_2} f(x)|^{x^j} = \frac{\partial}{\partial x_2} f(x)|^{x^i} + \frac{\partial}{\partial^2 x_2} f(x)|^{x^i}(x_2^j - x_2^i) + \frac{\partial}{\partial x_2 \partial x_1} f(x)|^{x^i}(x_1^j - x_1^i)$$

This is visualized in Figure 5.6, where $\Delta x_1 = (x_1^j - x_1^i)$ and $\Delta x_1 = (x_1^j - x_1^i)$.
From the previous step, we can directly evaluate the gradient terms

$$\frac{\partial}{\partial x_r} f(x)|^{x^k}, r \in 1, 2, k \in i, j$$

and wish to solve for the second-order terms, which will become the entries in the Hessian matrix. However, we have n^2 unknown second-order terms but only n equations, making this system underdefined. In order to have a fully-defined system, we need another set of difference equations, found by considering the difference with respect to another point k. Using the notation $(x_1)|_j^k = x_1^k - x_1^j$ and $\frac{\partial}{\partial x_1} f(x)|_j^k = \frac{\partial}{\partial x_1} f(x)|^k - \frac{\partial}{\partial x_1} f(x)|^j$ we can now rearrange the problem into the matrix equation

$$\overrightarrow{G} = D\overrightarrow{H}$$

or for the 2-D case:

$$\begin{bmatrix} \frac{\partial}{\partial x_1} f(x)|_j^k \\ \frac{\partial}{\partial x_2} f(x)|_j^k \\ \frac{\partial}{\partial x_1} f(x)|_i^k \\ \frac{\partial}{\partial x_2} f(x)|_i^k \end{bmatrix} = \begin{bmatrix} (x_1)|_j^k & (x_2)|_j^k & 0 & 0 \\ 0 & 0 & (x_1)|_j^k & (x_2)|_j^k \\ (x_1)|_i^k & (x_2)|_i^k & 0 & 0 \\ 0 & 0 & (x_1)|_i^k & (x_2)|_i^k \end{bmatrix} \begin{bmatrix} \frac{\partial}{\partial^2 x_1} f(x) \\ \frac{\partial}{\partial x_1 \partial x_2} f(x) \\ \frac{\partial}{\partial x_2 \partial x_1} f(x) \\ \frac{\partial}{\partial^2 x_2} f(x) \end{bmatrix}$$

In this form, where $\overrightarrow{G}$ collects the changes in gradient evaluations induced by each difference equation, the matrix D arranges the difference in coordinates at the evaluated iterates, and $\overrightarrow{H}$ unstacks the entries of the Hessian matrix H.

The entries of the Hessian matrix can then be found as $\overrightarrow{H} = D^{-1}\overrightarrow{G}$ or for the 2-D case:

$$\begin{bmatrix} \frac{\partial}{\partial^2 x_1} f(x) \\ \frac{\partial}{\partial x_1 \partial x_2} f(x) \\ \frac{\partial}{\partial x_2 \partial x_1} f(x) \\ \frac{\partial}{\partial^2 x_2} f(x) \end{bmatrix} = \begin{bmatrix} (x_1)|_j^k & (x_2)|_j^k & 0 & 0 \\ 0 & 0 & (x_1)|_j^k & (x_2)|_j^k \\ (x_1)|_i^k & (x_2)|_i^k & 0 & 0 \\ 0 & 0 & (x_1)|_i^k & (x_2)|_i^k \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial}{\partial x_1} f(x)|_j^k \\ \frac{\partial}{\partial x_2} f(x)|_j^k \\ \frac{\partial}{\partial x_1} f(x)|_i^k \\ \frac{\partial}{\partial x_2} f(x)|_i^k \end{bmatrix}$$

Expanding to n dimensions

Generalized to $r, s \in l$ dimensions (where $l \leq n$), this becomes:

$$\frac{\partial}{\partial x_r} f(x)|^{x^j} = \frac{\partial}{\partial x_r} f(x)|^{x^i} + \sum_{s=1}^{l} \frac{\partial}{\partial x_r \partial x_s} f(x)|^{x^i} (x_s^j - x_s^i)$$

As there are n dimensions in x, there are a total of n^2 unknown terms in the Hessian [2] . As each new point x^j has n dimensions, it adds a new set of n difference equations. In order to fully define the set of equations for computing the Hessian, we need to collect a total of $n+1$ linearly independent points (reference point plus differences with n points) to solve for the n^2 terms in H. As computing the gradient utilizes values from two iterates, this means that for a problem with n dimensions, the detection algorithm can be conducted on iterations $n+2$ and beyond.

In the following, we index dimensions by $1, 2, \ldots, n$ and points by $a, b, \ldots, k$ where k is used as the reference point. The vector $\overrightarrow{G}$ is thus composed by stacking the gradient evaluations:

$$\overrightarrow{G} = \begin{bmatrix} \frac{\partial}{\partial x_1} f(x)|_a^k \\ \frac{\partial}{\partial x_2} f(x)|_a^k \\ \vdots \\ \frac{\partial}{\partial x_n} f(x)|_a^k \\ \vdots \\ \frac{\partial}{\partial x_1} f(x)|_{k-1}^k \\ \frac{\partial}{\partial x_2} f(x)|_{k-1}^k \\ \vdots \\ \frac{\partial}{\partial x_n} f(x)|_{k-1}^k \end{bmatrix}$$

[2] or $n(n+1)/2$ if we utilize the fact that the Hessian is symmetric, i.e. that $\frac{\partial}{\partial x_1 \partial x_2} = \frac{\partial}{\partial x_2 \partial x_1}$. For simplicity, we will not take advantage of symmetry in this manuscript, though this symmetry can be utilized to accelerate the solution time of the algorithm.

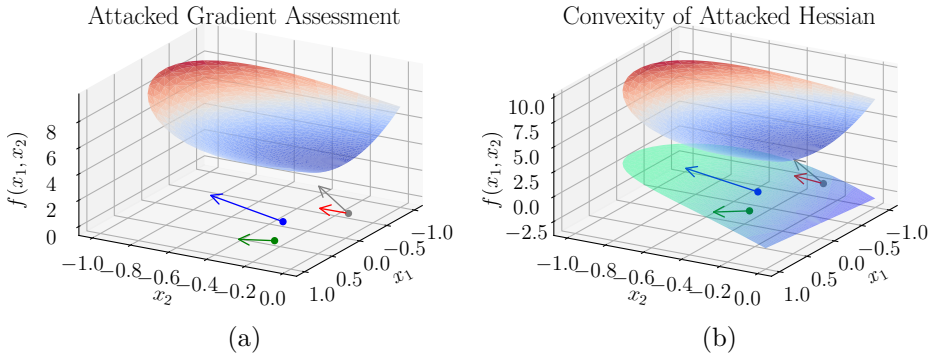


Figure 5.7: The introduction of noise shifts the gradients computed using the above algorithm, shown here as shifting the unattacked (gray) vector to a new (attacked, red) position. It can be readily seen that a sufficient displacement will cause the estimated curvature (Hessian) to become nonconvex, as shown by the inferred surface represented in Subfigure (b). To avoid detection, an attacker would thus need to use a very small injected signal, but this would not be sufficient to stop convergence and the attack would fail.

The matrix D is composed of a stack of n block-diagonal matrices, each of which is the Kronecker product of the n-dimensional identity matrix and a set of coordinate differences from x-iterates. Each block of this is $n \times n^2$, producing $D \in \mathbb{R}^{n^2 \times n^2}$:

$$D = \begin{bmatrix} I_n \otimes (x^k - x^a)^T \\ I_n \otimes (x^k - x^b)^T \\ \vdots \\ I_n \otimes (x^k - x^{k-1})^T \end{bmatrix}$$

We solve for the Hessian elements by computing $\overrightarrow{H} = D^{-1}\overrightarrow{G}$ as before.

Assess Convexity

For $f(x)$ to be convex, the Hessian must be positive semi-definite, i.e. $\lambda_{\min}(H) \geq 0$ where $\lambda_{\min}(H)$ is the least eigenvalue of H.

After solving for a local approximation of the Hessian as above, the eigenvalues of the Hessian are evaluated. If the Hessian is not found to be positive semi-definite, the x-actor must be injecting noise into his updates, as shown in Figure 5.7.

Misclassifications

As attack detection can be viewed as a classification problem, it may be subject to two possible errors, depicted in Figure 5.8.

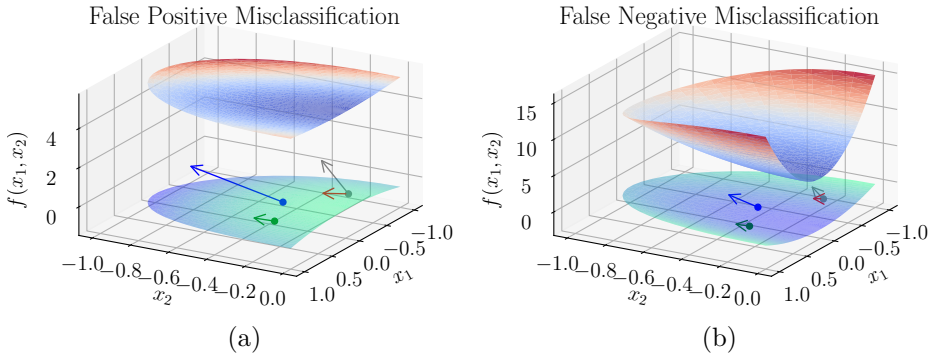


Figure 5.8: Depiction of misclassification errors in the noise detection algorithm. In weakly convex regions of a function, numeric conditioning errors can lead to false positives (subfigure a), whereas problems which have a very strong gradient may converge before the noise detection algorithm is able to identify the injection of noise into the algorithm (subfigure b). In both cases, the algorithm for selecting points used to assess the gradient can lead to more accurate assessment of the Hessian, reducing misclassification errors.

- False Positives (Type 1 errors): These occur when an attack is detected, even though no attack took place. This can occur when numeric issues create a slight local non-convexity during solution iterations, even though the underlying objective function is in fact convex. These numeric issues can result from the x-update, y-update, gradient calculation, or Hessian calculation, but are aggravated when the algorithm is close to the objective function and the system of equations used to generate the Hessian is poorly conditioned.

- False Negatives (Type 2 errors): In this type of error the algorithm fails to detect an attack, because the magnitude of the attack was not sufficient to create a non-convexity at the tested points. A small number of false positives are unavoidable: some attacks are not able to prevent convergence within a small number of iterations, and the attack detection algorithms may never be able to collect enough points to detect an attack. For problems that converge with a moderate number of iterations under attack, there may not be a combination of iterate points which result in a nonconvex Hessian. For other cases, lowering the false negative rate requires testing a variety of point combinations in order to maximize the likelihood of finding a point where the finite-difference Hessian estimate is nonconvex.

5.7 Simulation and Results

To verify the effectiveness of this problem under a variety of conditions, we simulate a large number of sample problems (both attacked and unattacked) and evaluate the effectiveness

of the attack detection algorithm.

We have chosen to consider quadratic programs (QPs): they are widely used in power systems research (e.g. optimal power flow [141], optimal dispatch [10,143]), subsume the set of linear programming problems, and yet offer many computational benefits (efficient off-the-shelf solvers, easy analytic solutions, easily visualized). In the future, we would like to extend this to other classes of convex problems (e.g. SOCP, SDP), but simulating hundreds of those problems would be computationally burdensome.

Problem: Random Quadratic Programs

We consider unconstrained quadratic programs of the form

$$\begin{aligned} \min x, z \quad & x^T P x + c^T x + z^T Q z + d^T z \\ \text{s.t.} \quad & Ax + Bz = c \end{aligned}$$

For simplicity, we examine problems where $c = 0$ and A and B have a single non-zero value per row, as described below. This can be thought of as consensus problems where some cost information is private.

Problem Generation

Problems are generated randomly as follows:

1. The size of the problem is determined by randomly choosing the magnitudes m, n, p. The user defines a maximum size `maxdim` and n and m are drawn as integers on the uniform distribution $[1, \texttt{maxdim}]$. The number of consensus constraints p is then randomly chosen by drawing an integer on the interval $[1, \min(m, n)]$.

2. The problem will then be characterized by:

   - Variables $x \in \mathbb{R}^n, z \in \mathbb{R}^m$,
   - Cost terms $P \in \mathbb{R}^{n \times n}, Q \in \mathbb{R}^{m \times m}, c \in \mathbb{R}^n, d \in \mathbb{R}^m$,
   - Constraint matrices $A \in \mathbb{R}^{p \times n}, B \in \mathbb{R}^{p \times m}$

3. Quadratic cost matrices P and Q are created by first randomly composing $L_P \in \mathbb{R}^{n \times n}, L_Q \in \mathbb{R}^{m \times m}$ with entries drawn randomly from the uniform distribution on $[-S, S]$. P and Q are then constructed to be symmetric positive semi-definite by construction as $P = (L_P)^T L_P, Q = (L_Q)^T L_Q$. To confirm that these are positive semi-definite, the eigenvalues are computed and checked to be greater than 0.

4. c and d are generated by randomly drawing values from $[-S^2, S^2]$, resulting in entries that are the same magnitude as those in P and Q.

5. A is composed as $A = [0 \in \mathbb{R}^{(n-p) \times (n-p)}, I \in \mathbb{R}^{p \times p}]$ and B is composed as $B = [I \in \mathbb{R}^{p \times p}, 0 \in \mathbb{R}^{(m-p) \times (m-p)}]$

Solution Method

We consider the case where consensus is desired between two actors who do not want to share information about their cost information P, c and Q, d. In this scenario, ADMM is a well-suited tool for achieving optimality without compromising privacy.

For this problem, the ADMM algorithm can be stated as:

$$\begin{aligned} x^{k+1} &= \text{argmin}_x \quad x^T P x + c^T x + \frac{\rho}{2}||Ax + Bz^k - c + u^k||_2^2 \\ z^{k+1} &= \text{argmin}_z \quad z^T Q z + d^T z + \frac{\rho}{2}||Bz + Ax^{k+1} - c + u^k||_2^2 \\ u^{k+1} &= u^k + Ax^{k+1} + Bx^{k+1} - c \end{aligned}$$

The x-update and z-update steps constitute unconstrained QPs, and can be solved analytically as described in Appendix A.1.These analytic update steps were used to avoid rounding issues resulting from using an iterative numeric solver.

Attack Implementation

Attack was simulated by multiplying the optimal x-update values by a vector with entries randomly selected from $+10\%, -10\%$ at each iteration. It was found that choosing uniformly distributed zero-centered noise was not sufficient to prevent convergence, as small-magnitude noise entires allow the algorithm to converge faster than the attacker is able to delay convergence.

A set of 10,000 QPs was simulated without attack, and were found to converge in a median of 63 iterations, with a mode of 16 iterations and a maximum value of 216 iterations. When simulated with the attack described above, convergence was significantly hindered on the same QPs: 86% were not able to converge in 300 iterations, and 87% do not converge in 500 iterations. The 14% of QPs which converge in less than 300 iterations were found to behave similarly to the unattacked QPs, likely due to a set of costs which are strongly convergent even in the presence of attack.

The distribution of the iterations until convergence for both the unattacked and attacked QPs is depicted in Figure 5.9.

Attack Detection Implementation

We implement the algorithm described above. As we initially consider $p \leq 2$, we need to complete at least 4 iterations in order to create two sets of difference equations.

To improve conditioning, for iteration i we consider the difference equations with i and the reference point, and compare with iteration 2 and iteration $\lfloor i/2 \rfloor$. Using the $i, i-1, i-2$ iterates was found to result in poor conditioning and a very high (17%) false positive rate, though a slightly lower false negative rate.

To ensure that we can solve for the values of the Hessian entries, we check that the difference vectors are not collinear.

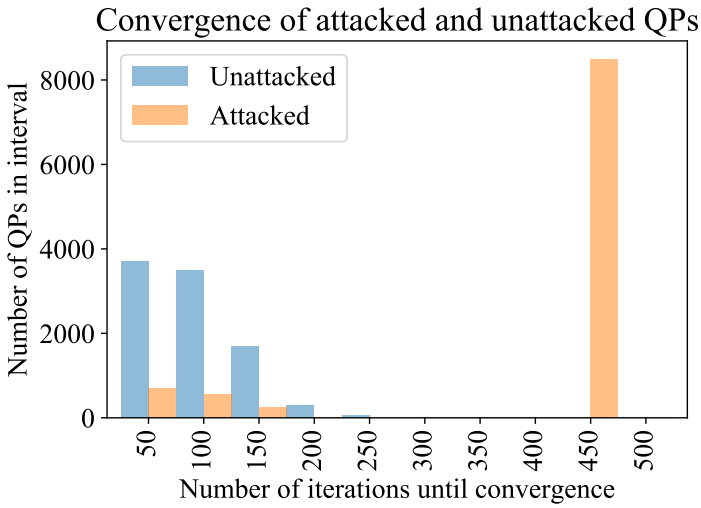


Figure 5.9: Convergence results for 10,000 QPs, simulated both with and without noise-injection attack. Without attack, all simulations converge in less than 300 iterations. Under attack, 86% of problems do not converge in 500 iterations (at which time calculation was stopped).

Table 5.1: Results for a simulation of attack detection. 500 quadratic programs were generated without attack, and false positives (upper right quadrant) were identified. An additional 500 quadratic programs were generated with simulated attack, and false negatives measured.

(a) Results, total number of simulations

	Attacked	Unattacked
Attack Detected	479	0
No Attack Detected	21	500
Total	*500*	*500*

(b) Results, simulation error rate

	Attacked	Unattacked
Attack Detected	0.958	0
No Attack Detected	0.042	1.0
Total	*1.0*	*1.0*

Results

We show results for 1000 simulated problems, of which 500 were attacked and 500 were unattacked. The results are presented as a *confusion matrix* in Table 5.1 which shows both accurate classifications and misclassifications. The number of problems of each type is shown in Table 5.1(a); the same results are shown in proportional form in 5.1(b).

We consider two types of errors: Type 1 (false positive, upper right quadrant) errors, and Type 2 (false negative, bottom leftquadrant) errors. The experimental results were compared when computing the local update with both numeric solvers and with an analytic solution; both approaches were found to give comparable results.

The choice of iterates used to compute the difference equations was found to have a significant impact on the results: as only $n+1$ points are needed but k iterates are available, there are $\frac{k!}{(n+1)!(k-n-1)!}$ possible choices for points to use, assuming that the most recent point is chosen as a base. A naive approach in which the $n+1$ most recent iterates were used was found to result in poor numeric conditioning and a large number of Type 1 errors; the results presented here select the $n+1$ iterates to be evenly spaced between iterate 2 and iterate k.

5.8 Limitations

This algorithm has been tested with constrained and unconstrained QPs of relatively small dimension, and in development was also tested with linear programs. We have not explored other problem types (e.g. second-order cone programs, semi-definite programs) where the Hessian may change over the iterates. We also have not expressly addressed the theoretical implications of using this on strongly-convex problems (e.g. problems which can be bounded below by a QP) compared with weakly-convex problems.

High-Dimensional Problems

We have previously assumed that $n+1$ iterates are needed to assess the validity of the x-update step. Where n is large, collecting sufficient iterates may be costly in time or computation resources. Further, the problem may converge before $n+1$ iterates are collected (not a problem if the goal is to prevent convergence-stopping attacks) and the ability to choose

We consider two scenarios: one where the full set of x-update variables are made public, and one where only the variables associated with the linking constraint are made public.

All x-variables are public

When all x-variables are made public, the gradient is constructed as:

$$\begin{aligned} A &\in \mathbb{R}^{p\times n} \\ x_{\text{pub}} &\in \mathbb{R}^{n} \\ (z^i - z^{i-1}) &\in \mathbb{R}^{m} \\ (y^i - \rho B(z^i - z^{i-1}) &\in \mathbb{R}^{p} \\ -A^T(y^i - \rho B(z^i - z^{i-1})) &\in \mathbb{R}^{n} \end{aligned}$$

The resulting gradient contains all x-dimensions, but will be rank deficient, having span $\mathbb{R}^p$ and nullspace $n-p$ In this case, each new point gives us n new dimensions, but we also seek to determine the Hessian in n dimensions and n^2 unknowns. We thus need to gather $n+1$ points.

Only linking dimensions are public

If only linking dimensions are public, the gradient is constructed as:

$$\begin{aligned} A &\in \mathbb{R}^{p\times p} \\ x_{\text{pub}} &\in \mathbb{R}^p \\ -A^T(y^i - \rho B(z^i - z^{i-1})) &\in \mathbb{R}^p \end{aligned}$$

In this case, the gradient estimate fully spans the linking dimensions, and each new point gives us p new equations. We seek the Hessian in p dimensions and p^2 unknowns, thus needing to gather $p+1$ datapoints.

Either way, we only are able to assess nonconvexity in the linking dimensions, and are unsure of activity in the other dimensions.

5.9 Extensions

Reducing Error Rates

As described above, the false positive and false negative rates can be reduced by tuning the algorithm:

- *Reducing False Positives:* As previously described, false positives result from poor numeric conditioning in computing the Hessian from the system of difference equations. Reducing these errors requires testing the condition number of the system of equations, and potentially also testing the condition number of the resulting Hessian estimate- both topics which have not been explored here.
- *Reducing False Negatives:* False negatives occur when the tested set of points do not indicate a non-convexity; this is most likely to occur when the magnitude of attack is small relative to the curvature of the x-update function. Reducing this error rate requires testing a variety of combinations of points in order to maximize the likelihood of finding a point where magnitude of the attack is large relative to the local curvature of the function.

Improving Algorithmic Efficiency

We have not explored opportunities to improve the efficiency of the algorithm, but highlight three opportunities here:

- **Exploiting symmetry in Hessian:** We have treated the Hessian as having n^2 unknowns, but as it is symmetric there are actually only $n(n+1)/2$ independent terms. Exploiting this structure would mean that only $\lceil (n+1)/2 \rceil + 1$ points are needed, but also requires pruning the equation set to avoid over-determining the system of equations.

- **Exploiting problem-specific structure of Hessian:** For quadratic programs, the Hessian matrix is constant throughout the problem, meaning that the difference equations do not need to be constructed with respect to a single point, but rather can be constructed using all combinations of points. In this case, we need to only use j iterates such that $\binom{j}{2} > \lceil (n+1)/2 \rceil + 1$, allowing the algorithm to be started faster, and allowing for more robust checking in the case of poor conditioning.

- **Choosing point set:** If the Hessian is computed multiple times per iteration to reduce error rates as described above, choosing the iterates wisely can improve numeric conditioning and reduce the need for extra computations. This has not been explored.

5.10 Extension: Fully-decentralized optimization

As the central information hub, the aggregator in a decentralized optimization problem must be trusted to provide fair updates to all the nodes. When the aggregator and the local nodes have the same incentives -e.g. they are all computing resources owned by the same entity- this is unlikely to present a conflict of interests.

However, when the aggregator has different incentives than the compute nodes, the local nodes may not trust the central aggregator- e.g. if the aggregator can extract additional profits by acting as a market maker. Further, aggregator-based decentralized computation is subject to several other weaknesses: it requires a high communication overhead by coordinating message-passing between all nodes (a significant hurdle for highly distributed systems like energy devices), and introduces a central point of failure in the case of communication/power outage or cyberattack.

To address these issues with aggregator-coordinated decentralized optimization, *fully-decentralized* algorithms have been developed which take advantage of problem structure to achieve consensus between nodes which directly share constraints, rather than passing information from all nodes to a centralized aggregator [144].

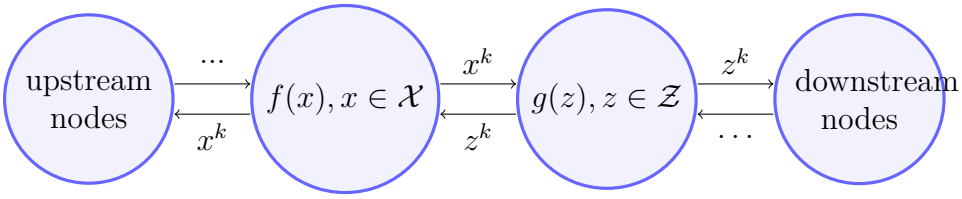


Figure 5.10: Fully-decentralized optimization problem structure, highlighting two nodes and noting how this can be indefinitely expanded with the addition of further upstream and downstream nodes. Each node holds private information on its own objective function, constraints, and dual variable (e.g. u_x^k), and accepts updates from its immediate neighbors.

Fully-Decentralized ADMM

As an example, the ADMM algorithm from above can be expressed in fully-decentralized form by introducing local copies u_x and u_z of the penalty variables:

$$x^{k+1} := \text{argmin}_{x\in\mathcal{X}}\ f(x) + \frac{\rho}{2}\|Ax + Bz^k - c + u_x^k\|_2^2 \tag{5.8}$$

$$z^{k+1} := \text{argmin}_{z\in\mathcal{Z}}\ g(z) + \frac{\rho}{2}\|Ax^{k+1} + Bz - c + u_z^k\|_2^2 \tag{5.9}$$

$$u_x^{k+1} := u_x^k + Ax^{k+1} + Bz^{k+1} - c \tag{5.10}$$

$$u_z^{k+1} := u_z^k + Ax^{k+1} + Bz^{k+1} - c \tag{5.11}$$

Under restrictions described in [144], this can be shown to have the same convergence and optimality guarantees as conventional aggregator-based systems.

In systems where nodes are very sparsely connected, this can produce a significant reduction in communication overhead- e.g. in power systems we may seek an optimum amongst millions of nodes, but each node is only connected to its parent and one or two downstream nodes. Rather than requiring that an aggregator handle millions of connections, nodes pass messages with their neighbors, ultimately bringing the full system into optimum. For examples of this approach applied to power systems, see e.g. [145–147].

5.11 Attack Vectors in Fully-Decentralized Optimization

As nodes in a fully-decentralized network are only connected with their neighbors, they must assess whether updates represent the true state of the network, or whether the neighbor has been compromised and the update is spurious. The following sections highlight the unique challenges of detecting, localizing, and mitigating attacks in a fully-decentralized system, as shown graphically in Figure 5.11.

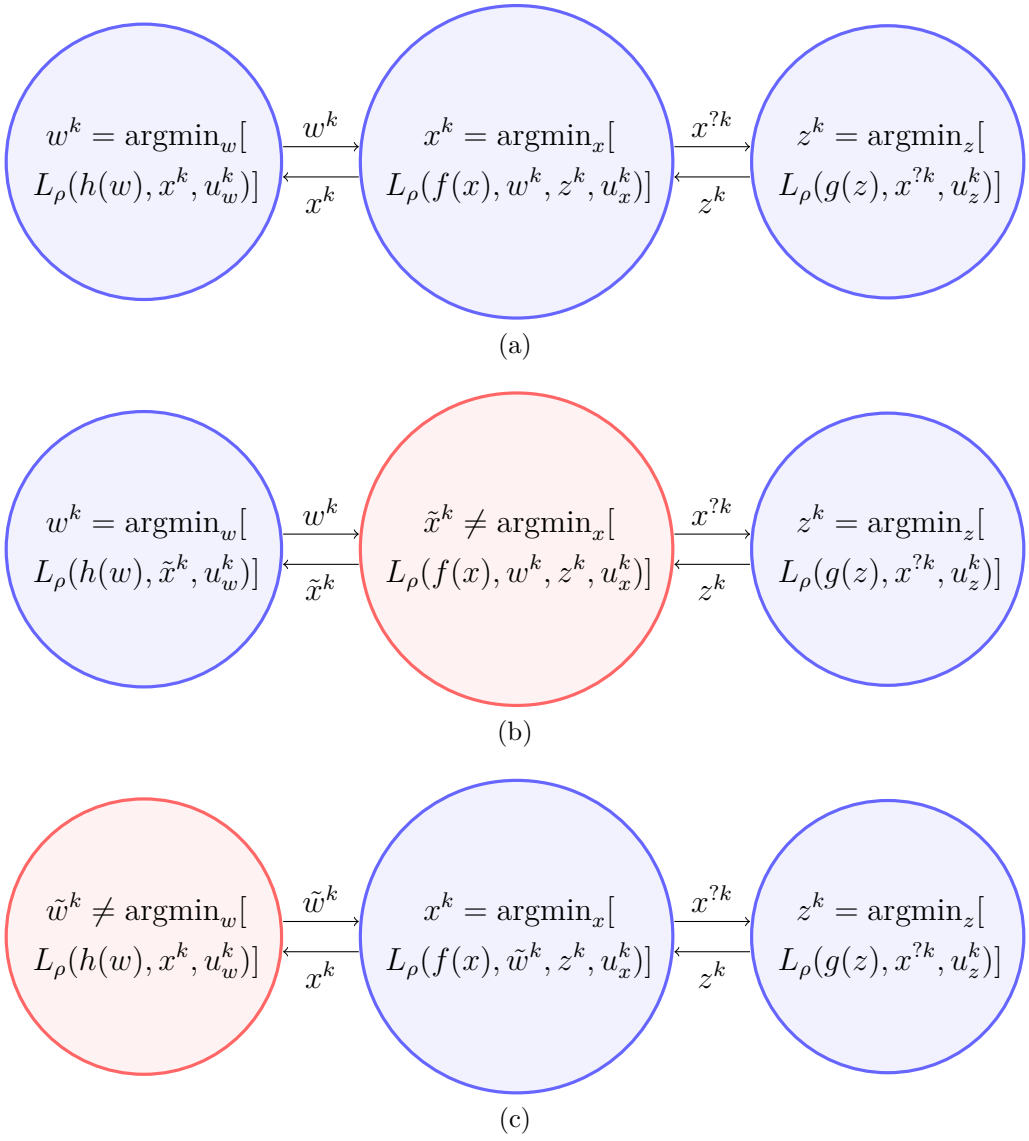


Figure 5.11: Potential attack scenarios in a fully-decentralized optimization scenario, where the z-update node is attempting to establish the veracity of a received update $x^{?k}$. Without knowing details of the private objective functions $f(x), h(w)$ and constraints $x \in \mathcal{X}, w \in \mathcal{W}$ it is difficult to detect and localize (or mitigate) an attack. The z-update node is not generally able to determine the difference between the unattacked scenario shown in (a), an attack by the immediate upstream neighbor $\tilde{x}^k$ as in (b), and an attack by a node further upstream such as $\tilde{w}^k$ as shown in (c).

Attack Vector: Private Infeasibility Attack

In Section 5.5, we highlighted how a malicious node may distort its private constraints to shift the equilibrium out of the operationally feasible region.

This attack is not identifiable in a fully-decentralized system, as a node can not in general discern between the update corrupted by its immediate neighbor

$$x^{?k} \stackrel{?}{=} \text{argmin}_{x\in\tilde{\mathcal{X}}} L_\rho(x, z^{k-1}, w^{k-1}, u_x^{k-1})$$

and a best response from a neighbor in reaction to a corrupted signal from the upstream $h(w)$ node:

$$x^{?k} \stackrel{?}{=} \text{argmin}_{x\in\mathcal{X}} L_\rho(x, z^{k-1}, \tilde{w}^{k-1}, u_x^{k-1})$$

Alternately, on seeing $x^{?k} \in \{\mathcal{X}_{\text{pub}}|z^{k-1}\}$ it is necessary to know w^k in order to assess whether the problem is truly feasible given the *full* state of the system, $x^{?k} \stackrel{?}{\in} \{\mathcal{X}|w^{k-1}, z^{k-1}\}$ (this can be extended to more complex system architectures with more upstream/downstream nodes).

However, in a fully-decentralized system, w is not generally visible to x and so it is not possible for the z-update node to assess whether $x^{?k} \in \{\mathcal{X}_{\text{pub}}|w^k\}$.

This means that detection, localization, and mitigation are all not possible in a conventional fully-decentralized system, as a global information layer is needed to assess the feasibility of the received updates, identify nodes which may be causing infeasibility, and construct a best response.

Attack vector: Infeasible Linking Constraint

Similarly, on receiving an update $x^{?k}$ which would violate the linking constraint $\{Ax^{?k}+Bz = c|z \in \mathcal{Z}\}$ the z-update node is not able to discern between an infeasible update created by a neighboring node:

$$x^{?k} \stackrel{?}{=} x^{\star k} + \varepsilon^k$$

and an infeasible signal resulting from a malicious upstream/downstream node:

$$x^{?k} \stackrel{?}{=} \text{argmin}_{x\in\mathcal{X}} L_\rho(x, z^{k-1}, \tilde{w}^{k-1}, u_x^{k-1}) = \text{argmin}_{x\in\mathcal{X}} L_\rho(z^{k-1}, z, w^{\star k-1} + \varepsilon, u_x^{k-1})$$

However, if information on public constraints $\mathcal{Z}_{\text{pub}}, \mathcal{X}_{\text{pub}}$ is knowable to all participants, these constraints can be integrated into the local optimization problems as additional (publicly known) constraints, converting the update step from

$$x^{k+1} = \text{argmin}_{x\in\mathcal{X}} L_\rho(x, z^k, u^k)$$

to

$$x^{k+1} = \text{argmin}_{x\in\mathcal{X}, x\in\{\mathcal{Z}_{\text{pub}}|A^Tx+B^Tz^k-c\}} L_\rho(x, z^k, u^k)$$

This would mean that *any* update which violates this constraint must be the result of a malicious node, and the problem can be localized.

Attack Mitigation

If a feasible solution exists, it must satisfy the constraints:

$$x^\star \in \mathcal{Z} \subset \mathcal{X}_{\text{pub}}$$
$$z^\star \in \mathcal{Z} \subset \mathcal{Z}_{\text{pub}}$$
$$A^T x^\star + B^T z^\star = c$$

Therefore, at each step we can project any iterate onto the feasible set described by the constraints above:

$$\hat{x}^k = \text{argmin}_{z \in \{\mathcal{Z} | A^T x + B^T z = c\}} ||x^{?k} - z||_2^2$$

In an unattacked case this projection simply accelerates convergence of the standard ADMM algorithm.

In the case of attack, projection creates a $\hat{x}^k \in \{\mathcal{Z} | A^T x + B^T z = c\}$ which is the *best response* to the attack $\tilde{x}^k$. While this best response will in general be suboptimal, it is feasible and will let the primal residual $r = Ax^k + Bz^k - c$ converge to zero even in the presence of attack.

Attack Vector: Zero-Mean Noise Injection

Because it is only dependent on local information (its own history and updates received from the neighbors), each node in a fully-decentralized network is able to carry out the detection strategy outlined in Section 5.6 for identifying the presence of a noise-injection attack.

However, in a fully-decentralized network the node would be challenged to identify whether the noise injection attack originates from an immediate neighbor, or from an upstream node.

Specifically, if $h(w)$ represents the problem solved by the upstream node, the z-update node cannot differentiate between $x^{?k} \stackrel{?}{=} x^{\star k} + \delta(k)$ and $x^{?k} \stackrel{?}{=} \text{argmin}_{x \in \mathcal{X}} L_\rho(x, z^{k-1}, \tilde{w}^{k-1}, u^{k-1}) = \text{argmin}_{x \in \mathcal{X}} L_\rho(x, z^{k-1}, w^{\star k-1} + \delta(k-1), u^{k-1})$

5.12 Summary of Security Challenges

Table 5.2(a) highlights the feasibility of achieving these goals in conventional aggregator-coordinated optimization, and Table 5.2(b) shows the same capabilities in a fully-decentralized system. We close by exploring approaches for reaching these goals in a fully-decentralized environment.

Architectures for Security

Although aggregator-based systems present some security and trust issues, the aggregator is able to take advantage of global information to check the validity of each node's updates,

Table 5.2: Security Issues in aggregator-coordinated and fully-decentralized systems

(a) Detection Feasibility on Aggregator-Coordinated System

	Detect	Localize	Mitigate
Private Infeasibility	X	X	X
Linking Infeasibility	X	X	
Noise Injection	X	X	

(b) Detection Feasibility on Fully-Decentralized System

	Detect	Localize	Mitigate
Private Infeasibility			
Linking Infeasibility	X	X	
Noise Injection	X		

enabling detection, localization, and mitigation strategies which are not available in the fully-decentralized system.

However, if the fully-decentralized system is augmented with a global information layer, the same security checks become possible in a fully-decentralized environment. In this section, we briefly discuss information architectures which could allow a fully-decentralized system to take advantage of the security benefits of an aggregator-coordinated system.

A number of different architectures might provide this security:

1. A centralized database held by a trusted authority, with security checks computed by each node

2. A fully-connected network with securely signed messages, in which each node maintains a database of message history

3. A partially-connected network with *bypass connections* to allow nodes to bypass suspect neighbors

4. A decentralized peer-to-peer database with message histories, with security checks computed by each node

5. A *blockchain* used as a decentralized database to store the message histories, with *smart contracts* used to compute security checks in a decentralized manner.

Option 1 presents trust and monopoly distortion issues described above, in addition to a high communication overhead. Option 2 requires high computational overhead and will fail if there are dropouts in communication with part of the network; Option 3 has gained some attention as a way to approach this while not requiring the same degree of communication redundancy. The decentralized database in Option 4 is appealing, but requires a method for reconciling different versions of the database which might be proposed by different neighbors.

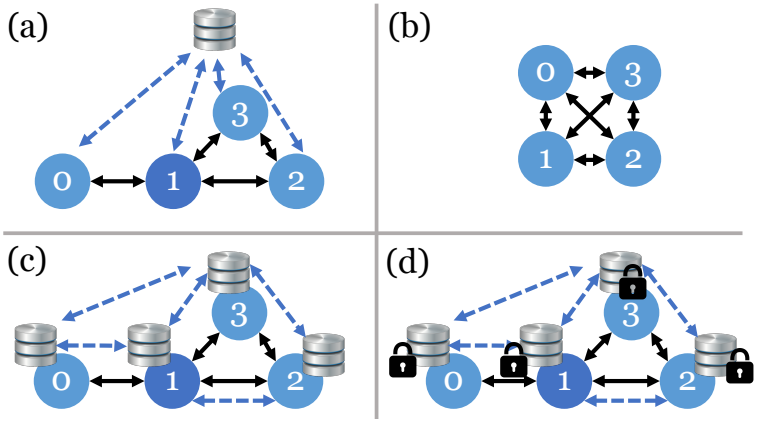


Figure 5.12: Potential architectures for allowing security checks for fully-decentralized optimization algorithms.

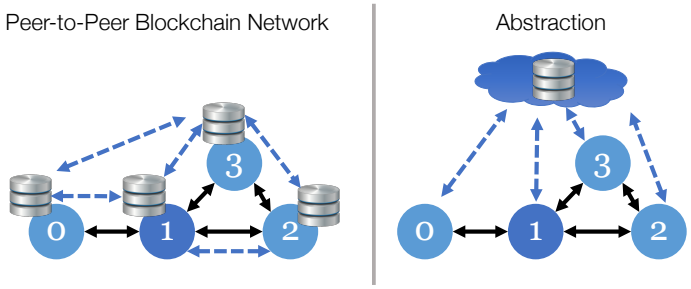


Figure 5.13: Illustration of how a blockchain-based network provides comparable security to an aggregator-coordinated architecture

Further, both Option 1 and Option 4 do not guarantee that each node conducts the same security checks.

By contrast, *blockchains* are decentralized databases which rely on securely signed messages and a consensus mechanism that simultaneously guarantees consistency across the network and provides secure timestamping. *Smart contracts* build on this architecture by guaranteeing the execution of simple computational functions as part of the consensus mechanism, thus guaranteeing consistent execution of security checks.

Figure 5.13 conceptually shows how a blockchain-based system enables the same type of security checks as are offered by an aggregator system, by providing a cryptographically secured global information layer with guaranteed execution of the security checks outlined above.

5.13 Potential Improvements

Fully-decentralized optimization models are appealing for their computational efficiency, privacy preservation, low communication overhead, and robustness to communication dropouts. However, this chapter shows that they are also vulnerable to attacks in which a node is compromised and deliberately broadcasts false information, leading the network away from the true optimum.

This chapter identifies techniques for securing these decentralized optimization problems, and highlights the need for a global information layer which can check that the updates provided by each node are feasible given the state of the network.

One technology for providing this global information layer is a *blockchain*, an emerging technology for decentralized computation and data storage which can achieve trustless concensus between nodes.

This *trustlessness* can not only improve security, but also allow decentralized optimization models to be used in scenarios where each party has incentives to cheat- for instance by allowing competitors to collaborate on a statistical estimation problem.

The next chapter applies these techniques along with the previous building blocks to create a model for secure fully-decentralized optimiation of energy dispatch on a microgrid. This scalable model computes energy prices endogenously, schedules flexible resources to minimize costs, and provides security checks through a blockchain and smart contract.

Chapter 6

Blockchains and Energy Control

The previous chapters have introduced flexible energy resources, optimization models, power flow models, and security paradigms in computation networks. This penultimate chapter ties together these different tools, exploring how large fleets of distributed energy resources can be controlled with a fully-decentralized optimization algorithm which is secured through smart contracts on a blockchain.

6.1 Introduction and Motivation

The energy production landscape is being reshaped by distributed energy resources (DERs) — photovoltaic panels, electric vehicles, smart appliances, and battery storage systems, which provide low-voltage energy services and are often remotely controllable as part of the Internet of Things. When used intelligently, these DERs can reduce cost, improve reliability, and integrate renewable resources in the electric grid — features which have led regulators to introduce policies promoting their adoption [148,149]. However, these new technologies raise two challenges: a technical challenge of coordinating large numbers of distributed devices, and an economic and regulatory challenge of incentivizing appropriate participation in energy services. In this paper, we present a framework for addressing both of these challenges.

Economic hurdle: Decentralized Energy Markets

Currently, payments for DER services must be negotiated with electric utilities, monopolies who may be invested in preserving conventional generation systems. As a result, the deployment of DERs has often been met with animosity by utilities, which may bar the participation of DERs or seek monopoly rents in return for access to the distribution infrastructure [150, 151].

Local distribution markets for energy services have been proposed as a means of efficiently incentivizing and dispatching DERs, much as is done at the transmission scale [152, 153]. However, such a local distribution market would need to address both the monopoly incentive

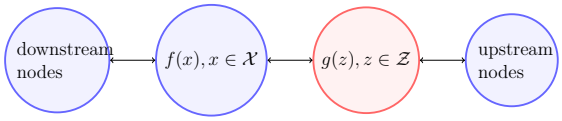


Figure 6.1: Sample problem structure, with arrows showing shared variables. Additional connections to upstream and downstream nodes may also be considered. The node computing $g(z), z \in \mathcal{Z}$ has been compromised.

issues highlighted above, and also the abuses of market power observed in wholesale energy markets [154, 155]. These issues would be particularly pertinent in microgrid operation, which may not benefit from the scrutiny given to a larger utility [156].

Technical challenge: Coordinating large decentralized fleets

Historically, electricity generation has been scheduled by central authorities or independent system operators, charged with ensuring that network constraints are met. However, these centralized optimization algorithms do not scale to large numbers generating units, and thus the deployment of large fleets of DERs has led to interest in *decentralized optimization* techniques. In these decentralized optimization models, the problem is broken into a set of small local problems which can be rapidly solved on distributed computing resources, with an *aggregator* bringing the local problems into consensus on a globally optimal solution.

These algorithms map readily to the architecture of the power network, where local parties have private information on consumption schedules and resource constraints, but need to achieve consensus on shared constraints such as power flow on the network.

Electricity distribution networks can be very large but are very sparsely connected: each node is connected to a single parent and a small number of children. In this environment, requiring all nodes to be connected to an aggregator creates a high communication burden; to reduce communication requirements *fully decentralized* models have been developed which remove the need for an aggregator by achieving consensus between neighboring nodes [147].

However, these approaches require *trust* between all nodes: without the presence of an aggregator, individual nodes must fully rely on their neighbors to pass correct information. In these fully-decentralized models, a compromised node is able to stall convergence on the network, or lead the system to an infeasible region, as discussed in [157].

Proposed solution: Smart Contracts

We leverage an emerging technology to address both the security and economic trust issues outlined above: blockchains and smart contracts. Despite extensive use in financial applications to achieve consensus between non-trusting parties [158], blockchains have seen limited deployment in the energy space [159] and have not been considered for coordinating DERs to manage network constraints [160].

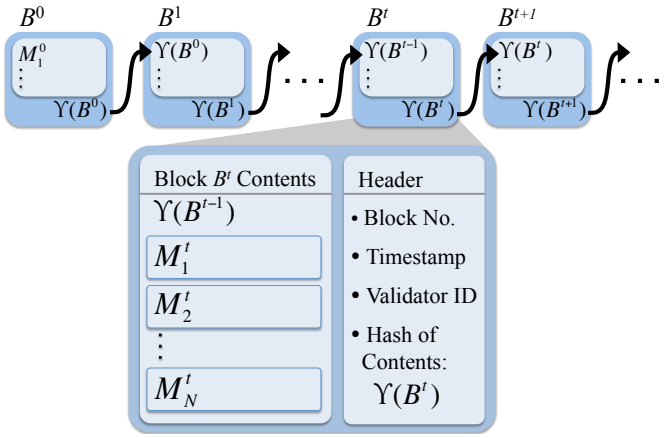


Figure 6.2: Symbolic representation of the data in a blockchain, showing blocks B^0 to B^{t+1} with detail of block B^t. Blocks are linked by their cryptographic hashes $\Upsilon(B^t)$, securing the contents from alteration and allowing transparent auditing of system history. Messages M_i^t contain information about changes to the system state, such as energy transfers or payments.

We examine how a blockchain architecture can be used to distribute the aggregator's role across all devices on a microgrid network. This integrated architecture is demonstrated on a blockchain platform controlling a microgrid simulation, and demonstrates how to address incentive issues while respecting operational constraints.

We structure the remainder of the paper as follows: Section 6.2 provides a brief overview of blockchains and smart contracts. Section 6.3 provides a survey of previous literature on dispatch of DERs in microgrids, decentralized optimization techniques, and blockchain use in energy applications. Section 6.4 presents the formulation of the optimal power flow problem with DERs and its ADMM equivalent, and Section 6.5 describes the algorithm for utilizing a blockchain for securing our decentralized problem. We present results from a simulation network, discuss limitations, and conclude by highlighting additional research opportunities.

6.2 Blockchains and smart contracts

Blockchains are an emerging technology for decentralized computation and data storage, secured by a combination of cryptographic signatures and a distributed consensus mechanism. Participants on the blockchain network are able to come to universal agreement on the system state σ^t at each time step t, even in the presence of cyberattacks, communication dropouts, and participants joining/departing the network. This is in stark contrast to conventional architectures where a central coordinator defines the state of the system, but may be subject to attack or malfeasance.

The general architecture of blockchains is described in [161] and illustrated in Fig. 6.2. Participants on the peer-to-peer network broadcast messages $M_i^t, i \in (1, \ldots, \mathrm{N_m})$. These messages contain commands which affect the state of the system (control actions, account

withdrawals, etc), and the feasibility of each message can be checked by each node using a validation function $\pi(\sigma^{t-1}, M_i^t)$.

Participants listen to the network and collect a set of messages into the contents of the next block B^t. A block header H is formed which contains the timestamp, a concise *cryptographic hash* $\Upsilon(B^{t-1})$ of the contents of the previous block, and the results of a verification test that is computationally or economically difficult to forge. The new block is broadcast to the network, where its validity is checked and nodes reach consensus on the updated state of the system $\sigma^t = \Pi(\sigma^{t-1}, B^t)$. The utility of blockchains can be significantly expanded when the state transition function $\Pi(\cdot)$ can execute computer code embedded in the transmissions M_i. These *smart contracts* can be transparently inspected and audited, and are guaranteed to be faithfully executed on the network.

Recursively linking the contents of blocks, verifying new blocks with peer-to-peer consensus, and using cryptographic signatures to verify communication are the pillars of blockchain architecture. Together, they provide an immutable and robust representation of system state- without requiring the intervention of a trusted central authority. While this architecture introduces some computational overhead, it offers immutability, transparency, and verifiability which can make the system well suited for coordination between parties who do not trust each other. We refer the reader to [8,158,160] for additional details on the security, architecture, and applications of blockchains and smart contracts.

6.3 Prior Literature

This work draws on three bodies of research: control of distributed energy resources, the economics and regulation of microgrids, and research on blockchains and smart contracts. We provide a brief summary of relevant literature from each domain.

DER Control

Microgrids are electricity networks which can be controlled autonomously, and may operate in both grid-connected and self-sufficient modes [162]. Without the benefit of a large balancing territory, loads and generation must be coordinated carefully and the role of DERs becomes particularly important. Surveys of approaches to microgrid management can be found in [163, 164].

Conventionally, generation resources have been centrally controlled by a utility or system operator — but these centralized approaches do not scale well to large numbers of DERs, and recent research has focused on decentralized algorithms with low computational overhead. Decentralized algorithms have been explored for coordinating electric vehicles [143, 165], smart inverters [166], and for fleets of diverse DERs [146, 147, 167].

Constraints on network voltage and power power flows can become significant at high DER penetrations, and decentralized models for power flow in distribution systems have been explored in [145, 168, 169]. As the underlying AC optimal power flow problem (OPF)

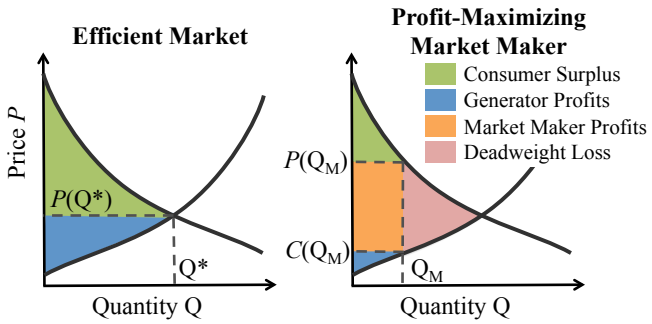


Figure 6.3: Comparison of an efficient market in which the optimal quantity $Q*$ is cleared at price $P(Q*)$, and a market operated by a monopoly who is able to charge separate prices for generation and consumption. In this model, the monopoly restricts output to Q_M, purchasing energy at $C(Q_M)$ and charging consumers $P(Q_M)$.

problem is nonconvex, each of these examines different assumptions or relaxations which grant computational tractability.

Microgrids and Monopoly Economics

In prior literature, DERs are compensated for providing energy services by an aggregator or a utility: a central authority who is trusted to act fairly in scheduling generators, satisfying loads, and rendering payments.

Like conventional electrical utilities, a microgrid operator faces a set of competing demands: minimizing consumer costs, investing in reliability and long-term capacity, and providing a return for shareholders [170]. Even without owning any assets, such a monopoly aggregator can have strong incentives to shift the market away from a cost-minimizing equilibrium and towards a profit-maximizing monopoly outcome, as shown in Figure 6.3. These conflicts of interest are typically controlled through regulatory intervention, where auditors scrutinize market outcomes and regulate customer fees [34].

However, when regulatory efforts are expensive, a small degree of market inefficiency may be less burdensome than regulatory costs [171]. This can create distrust between the microgrid operator and producers/consumers, who cannot assess whether their bills reflect monopoly profits or justified costs [148, 156].

This trust issue is already visible in the integration of rooftop photovoltaic systems in distribution networks [150, 151], and can be expected to be a greater problem in microgrids if regulatory scrutiny cannot be efficiently implemented for small systems — for example, if regulation has a high fixed cost (such as for retaining auditors) [171].

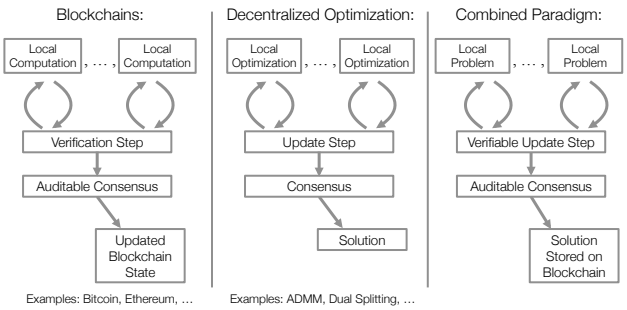


Figure 6.4: Conceptual models of the decentralized optimization involved in blockchain consensus networks, decentralized optimization problems, and our (novel) combined paradigm.

Blockchains and Energy

Blockchain research is still a new field, with most existing work focused on security and scalability [158, 160] and few applications for controlling physical devices [8, 172]. Although blockchains rely on a distributed consensus mechanism to provide security, the parallels with decentralized consensus algorithms in engineering control and optimization research have not yet been explored.

While blockchains have been discussed for use in coordinating DERs in transactive energy markets [5, 173], these works have not considered physical constraints on DER operation – instead treating DERs as idealized financial assets [174–176]. In reality, any coordination system must consider the DER's own constraints as well as the constraints of the distribution network. Prior literature has not considered methods for addressing these constraints in blockchain applications.

Novel Contributions

With this background, we see blockchains and smart contracts holding unexplored potential for eliminating trust issues with microgrid operators, and as a natural platform for coordinating the decentralized optimization schemes described above. The following contributions extend prior literature, creating the novel optimization paradigm found in Fig 6.4:

- Distributed optimal power flow algorithm with batteries, shapable loads, and deferrable loads
- Recovery of distributed locational marginal prices from a decentralized OPF problem
- Use of a blockchain for coordinating devices with operational constraints
- Use of a blockchain to facilitate the aggregator step of a decentralized optimization algorithm

6.4 Optimal Dispatch Formulation

We consider a microgrid with a dispatchable central generator, uncontrolled plug loads, non-dispatchable renewable energy resources, shapable loads (e.g. electric vehicles), deferrable loads (e.g. appliances), and batteries. We consider a day-ahead scheduling problem, with the objective of minimizing cost of energy provision subject to the operational constraints of the DERs and of the distribution network.

Note that in formulating the equations our variables are italicized, constants are non-italicized, and sets are denoted by calligraphy.

Network Model

The distribution network is modeled as an undirected radial graph $\mathcal{G}(\mathcal{N}, \mathcal{E})$, consisting of a set of nodes $\mathcal{N}$ and a set of distribution lines (a.k.a. edges) $\mathcal{E}$ connecting these nodes. Using the notation described in [177], we index the nodes in $\mathcal{N}$ by $i = 0, 1, \ldots, n$, where node 0 represents the root node (substation) and other nodes in $\mathcal{N}$ represent branch nodes. We also denote a line in $\mathcal{E}$ by the pair (i, j) of nodes it connects where j is closer to the feeder 0. We call j the parent of i, denoted by $\pi(i)$, and call i the child of j. Denote the child set of j as $\delta(j) := \{i : (i, j) \in \mathcal{E}\}$. Thus a link (i, j) can be denoted as $(i, \pi(i))$.

For each line $(i, \pi(i)) \in \mathcal{E}$, let $\mathrm{z}_i = \mathrm{r}_i + \mathbf{i}\mathrm{x}_i$ be the impedance of the line, let I_i be the complex current flowing from nodes i to $\pi(i)$, and $S_i = P_i + \mathbf{i}Q_i$ be the complex power flowing from nodes i to $\pi(i)$. On each node $i \in \mathcal{N}$, let V_i be the complex voltage, and $s_i = p_i + \mathbf{i}q_i$ be the net complex power injection. We assume the complex voltage V_0 at the substation node is given and fixed. We define $l_i := |I_i|^2$, $v_i := |V_i|^2$.

We model power flow on the network using the branch flow equations, first proposed in [178], which reflect a balanced single-phase radial network.

$$p_i = P_i - \sum_{k \in \delta(i)} P_k + \mathrm{r}_i l_i, \quad i = 0, \ldots, n \tag{6.1a}$$

$$q_i = Q_i - \sum_{k \in \delta(i)} Q_k + \mathrm{x}_i l_i, \quad i = 0, \ldots, n \tag{6.1b}$$

$$v_i = v_{\pi(i)} + 2(\mathrm{r}_i P_i + \mathrm{x}_i Q_i) - (\mathrm{r}_i^2 + \mathrm{x}_i^2) l_i, i = 1, \ldots, n \tag{6.1c}$$

$$l_i = \frac{P_i^2 + Q_i^2}{v_i}, \quad i = 1, \ldots, n \tag{6.1d}$$

where $S_0 = 0 + \mathbf{i}0$ at the slack bus. Equations in (6.1) define a system in the variables $(P, Q, l, v) := (P_i, Q_i, l_i, v_i, \ \forall \ \mathrm{i} \in \mathcal{N})$, which do not include phase angles of voltages and currents. Given (P, Q, l, v), phase angles can be uniquely determined for radial networks [179].

The final equation (6.1d) forms a non-convex set. It is relaxed to an inequality, which

yields a second-order cone constraint:

$$l_i \geq \frac{P_i^2 + Q_i^2}{v_i} \quad \iff \quad \left\| \begin{matrix} 2P_i \\ 2Q_i \\ l_i - v_i \end{matrix} \right\|_2 \leq l_i + v_i \tag{6.2}$$

In addition to the power flow equations, we also consider constraints on voltage magnitude on the network. These typically bound voltage within $\pm 5\%$ of a nominal voltage.

$$\underline{\mathrm{v}}_i \leq v_i \leq \overline{\mathrm{v}}_i, \quad i = 1, \ldots, n \tag{6.3}$$

Controllable DERs

We consider a set of energy resources with complex injections/withdrawals s placed at nodes i throughout the microgrid network, denoted as follows:

s_i^g	Dispatchable generators	s_i^u	Uncontrollable loads
s_i^r	Renewable generators	s_i^d	Deferrable loads
s_i^b	Stationary batteries	s_i^s	Shapable loads

The net complex injection at a node i in period t is

$$s_i(t) = s_i^g(t) - s_i^l(t), \quad i = 0, \ldots, n \tag{6.4}$$

where

$$s_i^l(t) = s_i^u(t) + s_i^d(t) + s_i^s(t) - s_i^b(t) - s_i^r(t) \tag{6.5}$$

Dispatchable generators (e.g. microturbines, diesel generators, fuel cells) are considered to have quadratic increasing cost, which may be time-varying:

$$C_{i,t}(s_i^g(t)) = \alpha_{i,t} s_i^g(t)^2 + \beta_{i,t} s_i^g(t) + \gamma_{i,t} \tag{6.6}$$

Power injection from renewable generators is considered to be deterministic and have no marginal cost $C_{i,t}(s_i^r(t)) = 0$. Power withdrawals due to uncontrollable loads (lights, plug loads) are considered deterministic, inflexible, and inelastic. We do not model thermostatically controlled loads or smart inverters, though those can be added to the formulation using the approaches in [180] and [145] respectively.

Stationary batteries are modeled as dispatchable loads which can be controlled to withdraw power ($s_i^b < 0$) or inject power ($s_i^b > 0$). We assume charging efficiency $\eta_{i,\mathrm{in}}$, and discharging efficiency $\eta_{i,\mathrm{out}}$. We assume that the battery should not undergo a net discharge

of more than ε over the course of the dispatch period.

$$\forall t = 1 \dots \mathrm{T}:$$
$$s_i^b(t) = d_i^b(t) - c_i^b(t) \tag{6.7a}$$
$$0 \leq c_i^b(t) \leq \mathrm{P}_{i,\mathrm{charge}}^b \tag{6.7b}$$
$$0 \leq d_i^b(t) \leq \mathrm{P}_{i,\mathrm{discharge}}^b \tag{6.7c}$$
$$\mathrm{E}_{b,\mathrm{min}} \leq E_b(t) \leq \mathrm{E}_{b,\mathrm{max}} \tag{6.7d}$$
$$E_i^b(t) = E_i^b(t-1) + c_i^b(t)\Delta\mathrm{t}\eta_{\mathrm{i,in}} - d_i^b(k)\Delta\mathrm{t}/\eta_{\mathrm{i,out}} \tag{6.7e}$$
$$(1+\varepsilon)E_i^b(1) \leq E_i^b(\mathrm{T}) \leq (1-\varepsilon)E_i^b(1) \tag{6.7f}$$

Shapable loads (e.g. electric vehicles with continuous charging levels, continuously variable fans) are modeled as having net energy demand $\mathrm{E}_{i,\mathrm{demand}}^s$, and must be charged between times $\mathrm{t}_{i,\mathrm{startby}}$ and $\mathrm{t}_{i,\mathrm{endby}}$:

$$\mathrm{P}_{i,\mathrm{min}}^s \leq s_s(t) \leq \mathrm{P}_{i,\mathrm{max}}^s \quad \forall t = 1 \dots \mathrm{T} \tag{6.8a}$$
$$\sum_{t=1}^{T} s_s(t) = \mathrm{E}_{i,\mathrm{demand}}^s \tag{6.8b}$$
$$s_i^s(t) = 0 \quad \forall t = 1, \dots, \mathrm{t}_{i,\mathrm{startby}} \tag{6.8c}$$
$$s_i^s(t) = 0 \quad \forall t = \mathrm{t}_{i,\mathrm{endby}}, \dots, \mathrm{T} \tag{6.8d}$$

Deferrable loads are considered to have some flexibility in their start time, but a defined load profile $l(\tau)\ \forall\ \tau = 1, \dots, L$ once started (e.g. appliances, manufacturing equipment). Following on the work in [146], we model the minimal starting time of the load as an arrival process $a(t)$, and the actual starting time as a departure process $d(t)$, where each of these variables takes the value 0 until the time of the request arrival/departure, at which point it takes the value 1. If the device can be started at most ζ time steps after the arrival request, we can formulate the constraints on our decision variable $d(t)$ as

$$\forall t = 1 \dots \mathrm{T}:$$
$$0 \leq d_i(t-1) \leq d_i(t) \leq a_i(t) \tag{6.9a}$$
$$a_i(t-\zeta) \leq d(t) \tag{6.9b}$$
$$d_i(t) \in (0, 1) \tag{6.9c}$$

Following [146] to formulate a matrix Φ which convolves the departure process $d(t)$ into a power consumption profile $s_d = \Phi d$, we can relax the binary constraint to allow scheduling to be expressed as a linear problem.

Optimal Power Flow

We consider the problem of maximizing social welfare in the network over a day, which amounts to scheduling the controllable loads to minimize generation cost, while respecting

network constraints. This problem is commonly known as *economic dispatch*, one of a family of optimal power flow (OPF) problems. It is formulated as follows:

$$\min \sum_{t=1}^{\mathrm{T}} \sum_{i=1}^{\mathrm{n}} C_{i,t}(s_i^g(t)) \tag{6.10a}$$

$$\text{s.t. } (6.1\text{a}), (6.1\text{b}), (6.1\text{c}), (6.2), (6.3), (6.4), \quad t = 1, \ldots, \mathrm{T} \tag{6.10b}$$

$$(6.7)_i, (6.8)_i, (6.9)_i, \quad i = 1, \ldots, \mathrm{n} \tag{6.10c}$$

$$\text{over } s_i^g(t) \in [\underline{s_i^g}, \overline{s_i^g}], i = 0, \ldots, \mathrm{n}, t = 1, \ldots, \mathrm{T}$$

$$(P_i, Q_i, l_i, v_i)(t), i = 1, \ldots, \mathrm{n}, t = 1, \ldots, \mathrm{T}$$

where constraints $(6.7)_i, (6.8)_i, (6.9)_i$, are specific to each node $i = 1, \ldots, n$, depending on the resources at that node.

In order to compensate the DER operators for their services and charge consumers for withdrawals, we want to compute nodal clearing prices, known as *distributed locational marginal prices* (DLMPs). The DLMP at a node represents the marginal cost to supply an additional unit of real power at that node. We denote the DLMP at node i as λ_i, and they can be found as the dual variables associated with the real power balance constraint (6.1a). As described in [177], the DLMP can be decomposed into contributions from energy, line losses, and voltage congestion.

Decomposition with ADMM

The Alternating Direction Method of Multipliers (ADMM) has gained popularity as a tool for decomposing difficult convex optimization problems into a set of simpler subproblems, coordinated through an aggregator step [4]. While convergence may be slow, the simplicity of the aggregator step and the guarantee of global optimality make the algorithm compelling for DER coordination. For examples of ADMM applications in various models of optimal dispatch problems, see [145, 147, 169, 181].

In the canonical ADMM problem, we consider a minimization problem with separable objectives and constraints in vectors x and z:

$$\begin{aligned} \min_{x,z} \quad & f(x) + g(z) \\ \text{s. to:} \quad & x \in \mathcal{K}_x, \ z \in \mathcal{K}_z \\ & Ax + Bz = c \end{aligned}$$

We can form the augmented Lagrangian:

$$L_\rho(x, z, \xi) := f(x) + g(z) + \xi^\top (Ax + Bz - c) + \frac{\rho}{2} \|Ax + Bz - c\|^2$$

This then decomposes into the general form of ADMM:

$$x^{k+1} = \arg\min_{x\in\mathcal{K}_x} L_\rho(x, z^k, \xi^k) \tag{6.11a}$$

$$z^{k+1} = \arg\min_{z\in\mathcal{K}_z} L_\rho(x^{k+1}, z, \xi^k) \tag{6.11b}$$

$$\xi^{k+1} = \xi^k + \rho(Ax^{k+1} + Bz^{k+1} - c) \tag{6.11c}$$

When decomposing a problem into subproblems for solution with ADMM, it is useful to think of x and z in the above as *local* and *global* variables respectively. Local variables only pertain to their respective subproblems, whereas global variables couple subproblems together and must be agreed upon at the global optimum, reaching a distributed consensus among subproblems. An intuitive way to formulate this is to give each subproblem its own copy of any coupling variables, and then try and make these copies agree.

The economic dispatch problem of (6.10) can be reformulated in this way by forming an individual subproblem at each node, whose solutions are made to coincide at the global optimum through copied local coupling variables. Each subproblem has its own copy of the relevant global coupling variable, and consensus on their value is achieved among subproblems through the ADMM algorithm. The subproblem of node i takes the following form, where for clarity we have omitted the time index from each nodal variable.

$$\min \sum_{t=1}^{T} C_{i,t}(s_i^g(t)) \tag{6.12a}$$

$$\text{s.t. } p_i = P_i - \sum_{k\in\delta(i)} P_k + \mathrm{r}_i l_i, \quad t = 1, \dots, T \tag{6.12b}$$

$$q_i = Q_i - \sum_{k\in\delta(i)} Q_k + \mathrm{x}_i l_i, \quad t = 1, \dots, T \tag{6.12c}$$

$$v_i = v_{\pi(i)} + 2(\mathrm{r}_i P_i + \mathrm{x}_i Q_i) - (\mathrm{r}_i^2 + \mathrm{x}_i^2) l_i, \ t = 1, \dots, T \tag{6.12d}$$

$$l_i \geq \frac{P_i^2 + Q_i^2}{v_i} \tag{6.12e}$$

$$(6.7)_i, (6.8)_i, (6.9)_i \tag{6.12f}$$

$$\text{over } s_i^g \in [\underline{s_i^g}, \overline{\mathrm{s}_i^g}], \mathrm{s}_i^b, s_i^d, s_i^s$$

$$(P_i, Q_i, l_i, v_i), (P_{\delta(i)}, Q_{\delta(i)}, v_{\pi(i)})$$

We first define a set of global variables $z := [P^\top, Q^\top, v^\top]^\top \in \mathbb{R}^{3n}$, a set of private local variable $x_i := [s_i^g, s_i^b, s_i^d, s_i^s, l_i]^\top$, and a set of coupling local variables $\tilde{x}_i = [P_i^\top, Q_i^\top, v_i^\top, P_{\delta(i)}^\top, Q_{\delta(i)}^\top, v_{\pi(i)}^\top]^\top$. We see that each subproblem i is coupled to other subproblems through the coupling local variables $\tilde{x}_i$, each of which is a selection of the components of the global variable z. Using notation from [4], the mapping from local variable indices into the global variable index can be written as $g = \mathcal{G}(i, j)$, which means that local variable component $(\tilde{x}_i)_j$ corresponds to global variable component z_k. Achieving consensus between the local variables and the global variable means that

$$(\tilde{x}_i)_j = z_{\mathcal{G}(i,j)}, \forall i, j \tag{6.13}$$

We can equivalently define a selection matrix B_i, such that, $\tilde{z}_i = B_i z$, and at the optimum

$$\tilde{x}_i - B_i z = \tilde{x}_i - \tilde{z}_i = 0 \tag{6.14}$$

At each iteration k, each node i, receives $\tilde{z}_i^k$ from the central aggregator, and solves

$$\begin{aligned} \min \ & \sum_{t=1}^{T} C_{i,t}(s_{i,t}) + \xi_i^{k\top}(\tilde{x}_i - \tilde{z}_i^k) + \frac{\rho}{2}\|\tilde{x}_i - \tilde{z}_i^k\|_2^2 \\ \text{s.t. } & (6.12) \\ \text{over } & x_i, \tilde{x}_i \end{aligned} \tag{6.15}$$

The node then sends its new $\tilde{x}_i^{k+1}$ to the central aggregator, who computes the following update for each individual global variable z_g^{k+1}

$$z_g^{k+1} := \frac{1}{\rho}\xi_i^k + \frac{1}{k_g}\sum_{\mathcal{G}(i,j)=g} (\tilde{x}_i^{k+1})_j \tag{6.16}$$

where k_g is the number of local variable entries that correspond to global variable entry z_g. The update can be thought of as taking the average of all local copies of the global variable. The central aggregator then updates ξ_i^k as

$$\xi_i^{k+1} = \ \xi_i^k + \rho(\tilde{x}_i^{k+1} - \tilde{z}_i^{k+1}) \tag{6.17}$$

We define the stopping criteria using the following residuals

$$r_i^k = \tilde{x}_i^k - \tilde{z}_i^k, \ s^k = z^k - z^{k-1} \tag{6.18}$$

Defining $r^k := [r_1^k, \ldots, r_n^k]$, the algorithm is determined to have converged when the both the following conditions are met

$$\|r^k\|_2 \leq \epsilon_{pri}, \ \|s^k\|_2 \leq \epsilon_{dual} \tag{6.19}$$

where ϵ_{pri}, ϵ_{dual} are suitably defined tolerances, and can be set using methods described in [4].

Fully-decentralized Decomposition

This approach can be extended using the fully-decentralized ADMM structure described in [144], which has been used in [146] and elsewhere for modeling a fully-decentralized ADMM model for control of energy resources in a distribution network.

As we consider radial distribution networks with no loops, we can create a bipartite labeling, where each node is an x-node surrounded by z-nodes, or vica versa. As discussed in [144] this is used for ordering our updates and creating a guarantee of convergence.

Rather than receiving a global variable estimate z from the aggregator, each x-node receives variable updates from its parent $\pi(i)$ and children $\delta(i)$, which we can collect into a variable representing all of its neighbors' estimates, z_i. At optimality, the local local shared variables must match the neighbors' estimates, creating the linking constraint $x_i = z_i$.

The full-decentralized ADMM algorithm can then proceed as a set of updates conducted at each of the x-nodes:

$$x_i^{k+1} = \arg\min_{x_i \in \mathcal{K}_{x_i}} L_\rho(x_i, z_i^k, u_i^k) \tag{6.20}$$

Note that this is the same as the x-update step in 6.11 and 6.15, but using the local coupling variables z_i and a local penalty term u_i computed by each node, rather than a global penalty computed by the aggregator.

These updated estimates for the x-nodes are then broadcast to neighbors, which will all be z-nodes thanks to the bipartite labeling. The z-update then proceeds as:

$$z_i^{k+1} = \arg\min_{z_i \in \mathcal{K}_{z_i}} L_\rho(x_i^{k+1}, z_i, u_i^k) \tag{6.21}$$

Finally, each node updates a local penalty u_i for deviations from the linking constraints.

$$u_i^{k+1} = u_i^k + \rho(x^{k+1} - z^{k+1}) \tag{6.22}$$

As in the aggregator-coordinated problem, the iterates are halted when the dual and primal residuals are below the acceptable threshold. In implementation, the fully-decentralized version was found converge significantly faster than the aggregator-secured system.

6.5 Blockchains and ADMM

We have formulated an optimal scheduling program for distributed energy resources through a decentralized algorithm. However, this only addresses part of the microgrid operation problem, and still has notable weaknesses:

- The aggregation step is not guaranteed against cyberattack or tampering by participants
- Individual DERs/consumers cannot verify that they are being paid/billed at fair prices
- Payments for actual generation/consumption will still be handled by a central utility

repeat
- **$\mathbf{P_i}$:** *Private Optimization, compute locally*
  - Gather private constraints
  - Compute $\tilde{x}_i$ and send to smart contract S_1
- **$\mathbf{S_1}$:** *Security Smart Contract, on blockchain*
  - Check $\tilde{x}_i$ for feasibility
  - If attack detected, flag node i & adjust $\tilde{x}_i$
  - Update u
  - **if** $\|r^k\|_2 \leq \epsilon_{pri},\ \|s^k\|_2 \leq \epsilon_{dual}$ **then**
    - Compute final schedule and clearing prices
    - Send schedule to S_2
  - **end**

until $\|r^k\|_2 \leq \epsilon_{pri},\ \|s^k\|_2 \leq \epsilon_{dual}$

$\mathbf{M_i}$: *Each Smart Meter*
- Record energy consumption
- Send time-stamped & signed consumption to S_2

...time progresses

$\mathbf{S_2}$: *Billing contract, on blockchain*
- Compare schedule from S_1 with meter readings
- Compute penalties, payments, and charges
- Transfer payments between accounts

Algorithm 1: Computational elements in the microgrid control system. Function $\mathbf{P_i}$ is executed locally by each device participating in the market. The results are passed to the smart contract $\mathbf{S_1}$, which serves as publicly verifiable ADMM aggregation step. $\mathbf{P_i}$ and $\mathbf{S_1}$ iterate back and forth until ADMM converges, at which point the schedule is saved to the billing smart contract $\mathbf{S_2}$. Smart meters send trusted meter readings to $\mathbf{S_2}$, which computes payments and automatically transfers funds from consumers to generators.

As an alternative, we propose to leverage the benefits of a blockchain architecture to create a fully peer-to-peer system which guarantees both operational feasibility and fair payments to all parties while taking full advantage of the decentralized structure of the problem, as discussed in [157].

As discussed in Section 6.2, blockchains provide a method for providing a transparent, trustless platform for data storage and computation. This makes a blockchain the perfect platform for conducting the aggregation step of ADMM, allowing all participants to audit the progress of the algorithm, the accuracy of the solution, and the veracity of their scheduled commitments. Further, ADMM is a natural fit for implementation on a blockchain, as it guarantees convergence yet has a *computationally cheap* aggregation step (minimizing the burden of verification).

Algorithm 1 provides an outline of the sequence of events in our proposed blockchain-

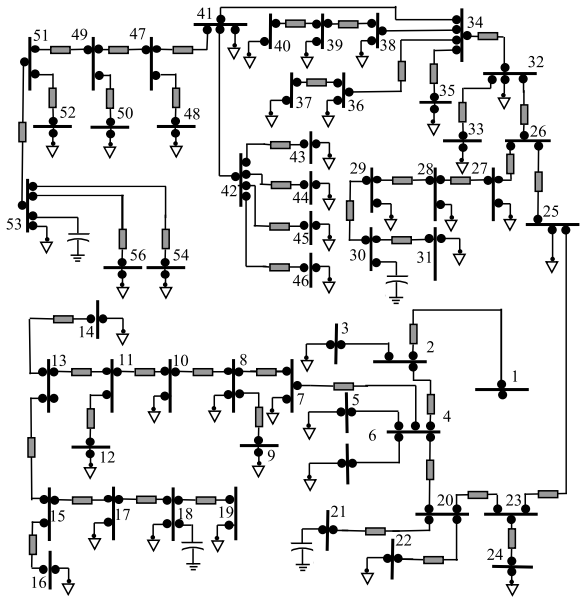


Figure 6.5: The 55-bus sample microgrid test feeder used in the simulation, with a microturbine placed at Bus 1 and DERs randomly distributed throughout the network.

based system. In it, we use the blockchain to (i) track the progress of variable updates, (ii) compute security checks on the proposed variable updates, (iii) store the resulting schedule, and (iv) compute payments and penalties for actual generation/consumption. Any participant can verify that the schedule maximizes social benefit while respecting network constraints, removing the possibility of monopolistic price manipulation.

This immutable record can also become the basis for reckoning payments if smart meters send consumption data to a billing contract S_2 which computes credits and debits for each node in the network and securely saves the updated account balance to the blockchain. Paired with a cryptocurrency as discussed in [158], this can form a complete payment system — removing the need for a utility or microgrid operator to handle scheduling and billing.

6.6 Implementation: Test network

We implement the proposed algorithm on a simulated Southern California Edison 55-bus test network shown in Fig. 6.5 with parameters described in [182]. Bus 1 is used as the reference bus, and is equipped with a large microturbine generator with quadratic cost function (this could also represent a connection to a utility grid). Each node has a deterministic load profile, created by adding a uniform random variable to the average uncontrollable load signal seen in Figure 6.6. We randomly placed solar arrays at 60% of the buses, and assume a deterministic solar generation profile. We place deferrable loads at 70% of the buses, with earliest start times randomly distributed between hours 7:00-11:00; these represent appliances and industrial equipment. Shapable loads are also randomly placed at 70% of

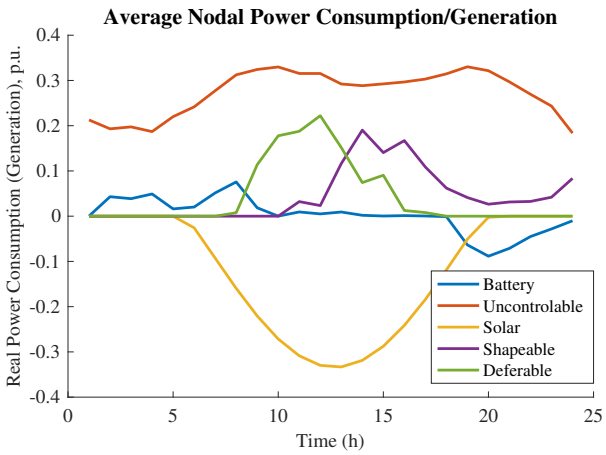


Figure 6.6: Schedule of commitments generated by the ADMM algorithm and stored to the smart contract. Positive values of power indicate real power consumption, and negative values indicate generation/injections.

the buses, with net energy demand generated from a uniform random variable that is up to 10 times the peak power consumption of the uncontrollable loads; these are intended to represent electric vehicle loads. The time constraints are randomly generated such that the shapable loads begin self-scheduling as early as 10:00, and can continue to draw power as late as hour 24:00. Batteries are placed at each bus, with a power capacity of 50% of peak controllable load at the bus, and with a 4 hour energy storage capacity.

We use a private Ethereum Homestead blockchain test network [183], and Python/CVXpy [184] to run the private optimization problems. Remote procedure calls through `EthJsonRpc` allow the Python scripts to communicate with the smart contracts.

Results and Discussion

The ADMM algorithm converged in 204 iterations, using $\rho = 100$, $\epsilon_{pri} = 10^{-3}$, $\epsilon_{dual} = 0.1$, with each iteration taking at most 1.2s to compute. The optimal cost of the distributed solution was 0.4% larger than the welfare-maximizing centralized OPF solution.

The average power consumption across all nodes is shown in Fig. 6.6. The power consumption profiles of individual nodes primarily differ in the temporal constraints, size of shapable load, and presence or absence of solar. The deferrable and shapable loads self-schedule to coincide with solar generation, while the battery charges and discharges to smooth net load.

The impacts of network topology can be seen in the voltage of each bus, shown in Figure 6.7. Since there are no current flow constraints, at optimality the upper voltage limit at the generator bus (#1) becomes the binding constraint; the voltages at each of the other buses decrease with distance from the feeder due to line effects (the critical link between bus 4 and 20 can be clearly seen). General trends in voltage over the course of the day are visible, with

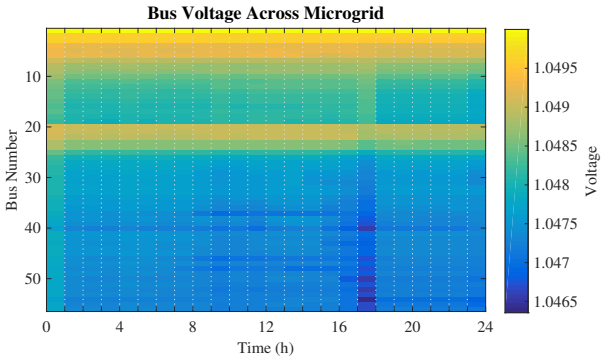


Figure 6.7: Voltage magnitude at each of the buses on the test network, for each hour in the simulation. Voltages vary based on local injections, and variations in time can be seen due to the impacts of local DER scheduling.

a significant drop in hour 18 when the setting sun and peaking uncontrollable load leads to a spike in net load throughout the network. Upon closer inspection, the impacts of DER scheduling are also visible at some buses (e.g. 38, 48.49) as appliances and EVs switch on and off.

The distributed marginal prices are not shown here for brevity, but can be easily calculated from the net load supplied by the central generator (other resources are inframarginal). We found very little variation in DLMPs between buses (variance of <1% of hourly DLMP), reflecting a lack of binding line constraints on this small network.

Security and attacks

To understand the potential vulnerabilities of a fully-decentralized ADMM algorithm, we consider two attack vectors, similar to those described in [157]:

- Convergence-stalling attack
- Operational infeasibility attack

To simulate an attack, we randomly designated 20% of the nodes as attackers, and examined system performance under the attack. We compare system performance in three scenarios: the original unattacked system, the attacked system without security checks, and the attacked system with security checks provided through a smart contract on the Ethereum blockchain as described above.

In the first attack mode, a compromised node broadcasts updates which are known to be infeasible for its neighbors, i.e. $z^k \in \{z \neq x \forall x \in \mathcal{X}\}$. By presenting these updates, the attacker prevents convergence of the primal residual, arbitrarily delaying the system and preventing a schedule from being issued by smart contract S_1. To simulate this, attacked nodes broadcast a voltage estimate of 0.90, outside of the allowable range of $[0.95, 1.05]$. In the unsecured system, the unattacked nodes are unable to achieve primal feasibility of the

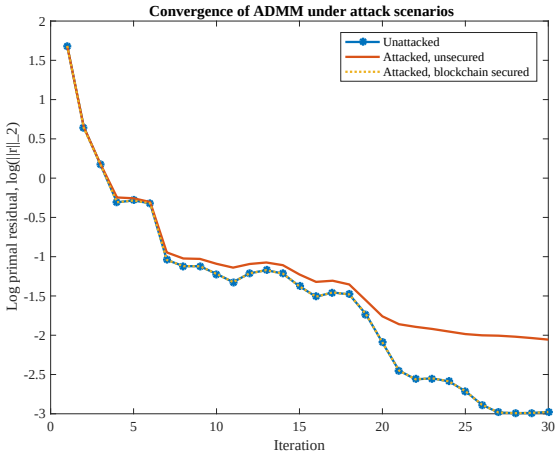


Figure 6.8: Convergence of the fully-decentralized ADMM algorithm under 3 modes of operation: normal operation (no compromised nodes), under the presence of attack from a node which is conducting a convergence-stalling attack, and under attack but secured with blockchain-based security checks.

coupling constraint $A_i\tilde{x}_i = B_i z_i$, and the primal residual remains outside of the convergence tolerance. This is shown in Figure 6.8, which shows the average value of the primal residual across all nodes in the network over the iterations of the ADMM protocol.

In the blockchain-secured system, the smart contract S_1 checks that all variable updates are feasible, and projects any infeasible (i.e. attacked) updates onto the known feasible set for the shared variables. This allows us to guarantee that the secured system can achieve convergence even in the face of attack. The blockchain-secured system is found to converge at the same rate as the original unattacked system, as the smart contracts lets us guarantee that all updates are feasible.

The second attack mode considers a compromised node which solves a *private problem* in which the private constraints have been modified to result in an optimum which is outside the original constraint set, i.e. $x_i \notin \mathcal{K}_{x_i}$. In a fully-decentralized system, this attack may not be detectable by individual nodes, as bounds on the feasible set $\mathcal{K}_{x_i}$ may be dependent on the upstream/downstream conditions of the node- in our case, the y are dependent on the upstream state of power flows in the system.

We simulate this second attack mode by introducing 'phantom' generation capacity at each of the attacked nodes. The attacked nodes thus replace Equation 6.4 with

$$s_i(t) = s_i^g(t) + s_i^p - s_i^l(t), \quad i = 0, \ldots, n \tag{6.23}$$

where s_i^p is the phantom generation, which is assumed to be three times the peak uncontrollable load, $s_i^p = k * \max(s_i^u)$, $k = 3$. Detecting this attack requires being able to compare the promised net injection $s_i(t)$ with known bounds on the node's injection capacity $\mathcal{S}_i$ and

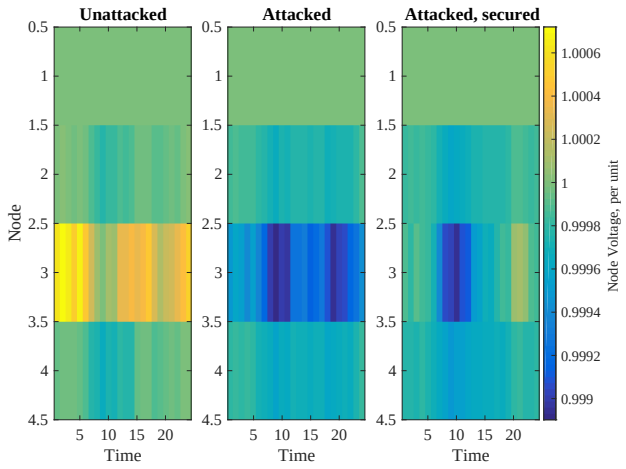


Figure 6.9: Scheduled voltage magnitude at each of the buses on the test network at each hour in the simulation, under three operating scenarios: normal operation (no nodes attacked), attacked by a malicious node which solves a privately infeasible problem, and a blockchain-secured algorithm. The security checks enabled by the blockchain system allow the network to reach a safe equilibrium in the presence of attack.

projecting the proposed update onto the feasible set $\mathcal{S}_i$ if an attack is detected. This is not feasible in a fully-decentralized system, as computing $s_i(t)$ requires sharing knowledge between all of a node's neighbors.

In the unsecured system, the attacked nodes then pass their shared variables to upstream and downstream nodes, leading the system to an operating schedule which is dependent on this phantom generation capacity. In operation, this could lead to violations of line flow or nodal voltage constraints, as the system struggles to make up for the missing generation capacity. We note that this is only one of many possible ways that the private constraints might be attacked- other attack modes might be to overstate/understate the power demand of deferrable and shapable loads or shift expected load times.

While convergence is not affected under this mode of attack, the operational performance will be impacted. We simulate this by taking the scheduled generation sent to S_2, and simulating performance of the system under this schedule. Imbalances from the schedule can result in nodal voltage violations, possible violation of line flow constraints (not considered here), or increased costs from importing additional power to the microgrid.

Figure 6.9 shows the nodal voltages which are expected in the unattacked, attacked, and the attacked but blockchain-secured scenarios. We see significant voltage deviations under the attacked scenario, pushing the system towards violation of the nodal voltage constraints. Under the blockchain-secured scenario, the system arrives at a schedule which mitigates these deviations, within the bounds possible given the limited public knowledge on the constraint set $\mathcal{S}_i$. For a more detailed discussion of the capabilities and weaknesses of the blockchain-based security system, see [157].

6.7 Limitations

The communication overhead required for ADMM and the verification delay required for a blockchain may limit this approach to use for day-ahead scheduling, while secondary and tertiary control is served by conventional approaches. We envision a blockchain-based economic scheduling layer (described here) operating in tandem with a real-time control layer operating at much faster time scales to address real-time imbalances.

As formulated, the local optimization problem holds the device constraints, thus preserving a degree of privacy for the energy consumers. However, the common variables include the expected energy schedules, and the blockchain publicly holds the commitments and delivered power information, effectively making smart meter data public knowledge. This may be a barrier for deployment in some applications, but we expect that in isolated microgrids and grid extension projects this condition would be acceptable. We are exploring the potential of using zero-knowledge proofs [185] as a method for creating data confidentiality while still preserving guarantees of convergence and optimality.

While the ADMM and D-ADMM algorithms offer guarantees of convergence, waiting for variable updates to be registered on the blockchain introduces an additional delay into each iteration. On the Ethereum network, default block formation time is set to be approximately 15s, creating a limit on the total speed in which an operating schedule can be formulated (100 iterations would take at least 25 minutes). The parameters of the blockchain can be adjusted to increase block creation rate [186], but this must be balanced with the latency in the rest of the network to prevent forks in the blockchain.

While this paper envisions each building's smart meter acting as a computational node in the blockchain network, alternate configurations may use devices, feeders, or substations as nodes. In each of these implementations, we assume that the network topology is fully known by all parties, and have not considered changes in line impedances (e.g. due to temperature changes) or in topology (e.g. due to outages).

This paper relies on previous work for proofs of the security, transparency, and robustness of blockchain-based systems (see e.g. [161,186]), and in future work we will explore the specific security concerns of the microgrid system outlined here.

6.8 Conclusions

We have shown how decentralized consensus techniques and blockchains can be used both to coordinate the scheduling of distributed energy resources on a microgrid, and to guarantee fair payments without requiring a utility or centralized microgrid aggregator. By using ADMM, we decompose our problem into a structure that naturally lends itself to a blockchain implementation, and show how blockchains and smart contracts can provide a natural solution for the trust, security, reliability, and immutability requirements of microgrid operation. We show the results are equivalent to a welfare-maximizing centralized dispatch with per-

fect insight into device constraints, yet avoid the risk of monopoly price manipulation and privacy concerns.

The proposed architecture can be improved with contributions from active areas of control research: addressing stochastic/uncertain data through model predictive control and robust optimization, examining resilience to network interruptions, utilizing fully distributed ADMM between nodes to reduce communication overhead, and developing fault detection algorithms to identify fraud and changes in system architecture.

While this is the first paper (to our knowledge) to examine the integration of blockchain with distributed optimization of energy systems, we expect that many other applications are possible, both within the energy sector and in other engineering realms. Blockchain's distributed consensus mechanism has proven itself in the finance world by guaranteeing robust, trustless, and transparent execution; we highlight similar benefits for controlling physical devices. We see blockchains and smart contracts as a key technology that enables distributed optimization amongst non-trusting entities, at all scales of operation.

6.9 Potential Improvements

While this demonstration culminates our research, it should not be viewed as the canonical or definitive example in this space. This new framework for blockchain-secured decentralized optimization can be applied to a wide variety of constrained optimization problems both in the energy management space and beyond, preserving the flexibility of the underlying convex optimization tools.

Many extensions to the model proposed here are possible, and could extend the variety of devices which can be integrated, the structure of the energy market and the network flow model, and improve convergence time. Additional research in the blockchain structure would be helpful to improve scalability to larger networks, create resilience to sharding and local blackouts, and reduce the network communication requirements and time to achieve consensus on the network state.

From discussions with utilities and industry leaders in the blockchain-energy space, we would expect that this approach will be most valuable in markets with high energy prices, high penetration of intermittent renewables and controllable energy loads, distribution constraints, and a restructured retail market for electricity. Wider adoption is not expected in the next 3-5 years.

Chapter 7

Conclusion

This dissertation has brought together tools from the fields of optimization theory, economics, power systems, and cybersecurity to develop a new framework for secure fully-decentralized computing in constrained systems. We have gradually introduced relevant tools over the preceding chapters to create a toolkit with which the reader can launch into additional research that can improve the operation of energy markets and particularly the integration of decentralized energy resources.

While we conclude our discussion with a blockchain-secured microgrid optimization model that ties together the different topics of the research, this should not be viewed as the canonical or definitive example in this space. This new framework for blockchain-secured decentralized optimization can be applied to a wide variety of constrained optimization problems both in the energy management space and beyond, preserving the flexibility of the underlying convex optimization tools.

Given the challenges of using smart contracts to verify that physical operations have been carried out as promised (e.g. that power was generated, and that hardware was not hacked), it is quite likely that this framework will see earlier deployment in machine learning or statistical estimation tasks between untrusting parties. However, by first proving that this system is feasible for systems with more complicated constraint sets, we hope to pave the path for future deployments in the energy space.

As blockchain technology evolves, we see great potential for unifying the decentralized consensus model used for blockchain security with the decentralized consensus models being explored in power system optimization. Additional research may focus on proving the scalability, convergence, and speed of these systems, particularly in an environment where full network information (e.g. network topology, circuit breaker status) is not readily obtained. This research should provide good fodder for future exploration.

Bibliography

[1] R. Poli, J. Kennedy, and T. Blackwell, "Particle swarm optimization An overview," *Swarm Intelligence*, vol. 1, pp. 33–57, 2007.

[2] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[3] L. El Ghaoui and G. Calafiore, *Optimization Models*. Cambridge University Press, 2014.

[4] S. Boyd, "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2010.

[5] J. Horta, D. Kofman, and D. Menga, "Novel paradigms for advanced distribution grid energy management," 2016.

[6] W. F. Samuelson and S. G. Marks, *Managerial Economics*. John Wiley & Sons, Inc., 7 ed., nov 2011.

[7] S. a. Gabriel, A. J. Conejo, D. Fuller, B. Hobbs, and C. Ruiz, *Complementarity Modeling in Energy Markets*. Spring International Series in Operations Research & Management Science, 2010.

[8] K. Christidis and M. Devetsikiotis, "Blockchains and Smart Contracts for the Internet of Things," *IEEE Access*, vol. 4, pp. 2292–2303, 2016.

[9] A. Von Meier, *Electric Power Systems: A Conceptual Introduction*. IEEE Press / Wiley-Interscience.

[10] R. Dobbe, D. Arnold, S. Liu, D. Callaway, and C. Tomlin, "Real-Time Distribution Grid State Estimation with Limited Sensors and Load Forecasting," 2016.

[11] California ISO, "What the duck curve tells us about managing a green grid," pp. 1–4, 2013.

[12] International Energy Agency, "The Power of Transformation: Wind, Sun and the Economics of Flexible Power Systems," 2014.

[13] I. Bolliger, E. Munsing, and J. Romankiewicz, "Evaluation of overgeneration in California's electricity market and discussion of potential solutions," *unpublished*, no. December, 2014.

[14] North Carolina Solar Center, "Renewable Portfolio Standard Policies," 2013.

[15] Energy and Environmental Economics (E3), "Investigating a Higher Renewables Portfolio Standard in California," Tech. Rep. January, Energy and Environmental Economics, Inc., San Francisco, CA, 2014.

[16] J. H. Williams, A. DeBenedictis, R. Ghanadan, A. Mahone, J. Moore, W. R. Morrow, S. Price, and M. S. Torn, "The technology path to deep greenhouse gas emissions cuts by 2050: the pivotal role of electricity.," *Science (New York, N.Y.)*, vol. 335, pp. 53–9, jan 2012.

[17] C. D. Scown, M. Taptich, A. Horvath, T. E. McKone, and W. W. Nazaroff, "Achieving deep cuts in the carbon intensity of U.S. Automobile transportation by 2050: Complementary roles for electricity and biofuels," *Environmental Science and Technology*, vol. 47, no. 16, pp. 9044–9052, 2013.

[18] S. Kiliccote, P. Sporborg, I. Sheikh, and M. A. Piette, "Integrating Renewable Resources in California and the Role of Automated Demand Response," *Lawrence Berkeley Natuional Lab Report*, vol. LBNL-4189E, no. November, 2010.

[19] J. Eyer and G. P. Corey, "Energy Storage for the Electricity Grid : Benefits and Market Potential Assessment Guide," *Sandia National Laboratories Report*, no. SAND2010-0815, 2010.

[20] California Public Utilities Commission, "Decision Adopting Energy Storage Procurement Framework and Design Program," *CPUC Decision*, vol. 13-10-040, 2013.

[21] J. Nelson, J. Johnston, A. Mileva, M. Fripp, I. Hoffman, A. Petros-Good, C. Blanco, and D. M. Kammen, "High-resolution modeling of the western North American power system demonstrates low-cost and low-carbon futures," *Energy Policy*, vol. 43, pp. 436–447, 2012.

[22] A. Mileva, J. H. Nelson, J. Johnston, and D. M. Kammen, "Sunshot solar power reduces costs and uncertainty in future low-carbon electricity systems," *Environmental Science and Technology*, vol. 47, no. 16, pp. 9053–9060, 2013.

[23] G. M. Morrison, S. Yeh, A. R. Eggert, C. Yang, J. H. Nelson, J. B. Greenblatt, R. Isaac, M. Z. Jacobson, J. Johnston, D. M. Kammen, A. Mileva, J. Moore, D. Roland-Holst, M. Wei, J. P. Weyant, J. H. Williams, R. Williams, and C. B. Zapata, "Comparison of low-carbon pathways for California," *Climatic Change*, 2015.

[24] K. Y. Chay and M. Greenstone, "Does Air Quality Matter? Evidence from the Housing Market," *Journal of Political Economy*, vol. 113, no. 227, pp. 376–424, 2005.

[25] D. Howarth and B. Monse, "Renewables Face Daytime Curtailments in California."

[26] L. Xu and D. Tretheway, "Flexible Ramping Products, Draft Final Proposal," *California ISO*, no. December, 2014.

[27] M. D. Tabone and D. S. Callaway, "Modeling Variability and Uncertainty of Photovoltaic Generation: A Hidden State Spatial Statistical Approach," *IEEE Transactions on Power Systems*, pp. 1–9, 2014.

[28] H. Trabish, "Hawaiian Electric's plan to end solar net metering, explained," jan 2015.

[29] Hawaiian Electric Companies, "Transient Over-Voltage and Frequency & Voltage Ride-Through Requirements for Inverter-Based Distributed Generation Projects," no. February, pp. 1–6, 2015.

[30] Department of Energy, "SunShot Initiative Technical Topics: Distribution," 2014.

[31] B. Chabot, "Personal Communication," *Pacific Gas & Electric Company*, no. May, 2015.

[32] J. Cart, "Saving desert tortoises is a costly hurdle for solar projects," mar 2012.

[33] E. Holden, "Strange Bedfellows: Environmental Groups, Transmission Developers Working Together on Renewable Energy Projects," apr 2014.

[34] R. Hirsh, *Power Loss: The Origins of Deregulation and Restructuring in the American Electric Utility System*. Cambridge, MA: MIT Press, 2000.

[35] S. Borenstein, J. Bushnell, and F. Wolak, "Measuring Market Inefficiencies in California ' s Restructured Wholesale Electricity Market," *American Economic Review*, vol. 92, no. 5, pp. 1376–1405, 2002.

[36] D. Glick, M. Lehrman, and O. Smith, "Rate Design for the Distribution Edge," *Rocky Mountain Institute*, no. August, 2014.

[37] R. Sioshansi, P. Denholm, and T. Jenkin, "Market and policy barriers to deployment of energy storage," *Economics of Energy & Environmental Policy*, vol. 1, pp. 1–14, apr 2012.

[38] P. Denholm, J. Jorgenson, T. Jenkin, D. Palchak, B. Kirby, and M. O. Malley, "The Value of Energy Storage for Grid Applications," *National Renewable Energy Laboratories report*, no. NREL/TP-6A20-58465, 2013.

[39] S. J. Kazempour, M. Hosseinpour, and M. P. Moghaddam, "Self-scheduling of a joint hydro and pumped-storage plants in energy, spinning reserve and regulation markets," in *2009 IEEE Power and Energy Society General Meeting, PES '09*, pp. 1–8, 2009.

[40] B. Kirby, "Co-optimizing energy and ancillary services from energy limited hydro and pumped storage plants," in *HydroVision*, pp. 1–11, 2012.

[41] B. Dunn, H. Kamath, and J.-M. Tarascon, "Electrical energy storage for the grid: a battery of choices," *Science*, vol. 334, no. 6058, pp. 928–935, 2011.

[42] R. Walawalkar, J. Apt, and R. Mancini, "Economics of electric energy storage for energy arbitrage and regulation in New York," *Energy Policy*, vol. 35, pp. 2558–2568, apr 2007.

[43] D. of Energy, "Grid Energy Storage," Tech. Rep. December, Department of Energy, 2013.

[44] Energy and Environmental Economics (E3), "Evaluation of Hawaii's Renewable Energy Policy and Procurement," no. January, 2014.

[45] M. Reguant and K. Ito, "Sequential Markets, Market Power and Arbitrage," *National Bureau of Economic Research*, vol. No. w20782, no. 7, pp. 1–54, 2014.

[46] G. Papaefthymiou, K. Grave, and K. Dragoon, "Flexibility options in electricity systems," *Ecofys*, no. POWDE14426.

[47] S. Schonung, "Energy Storage Systems Cost Update," *Sandia National Laboratories Report*, no. SAND2011-2730, 2011.

[48] Maxim Integrated, "Linear regulators in portable applications," *Application Note*, p. number 751, Aug. 2012.

[49] R. J. M. Vullers, R. van Schaijk, I. Doms, C. Van Hoof, and R. Mertens, "Micropower energy harvesting," *Solid-State Electronics*, vol. 53, no. 7, pp. 684–693, 2009.

[50] A. Harb, "Energy harvesting: State-of-the-art," *Renewable Energy*, vol. 36, no. 10, pp. 2641–2654, 2011.

[51] X. Hu, N. Murgovski, L. M. Johannesson, and B. Egardt, "Comparison of three electrochemical energy buffers applied to a hybrid bus powertrain with simultaneous optimal sizing and energy management," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 3, pp. 1193–1205, 2014.

[52] B. Kim, R. Winslow, I. Lin, K. Gururangan, J. Evans, and P. Wright, "Layer-by-layer fully printed Zn-MnO2 batteries with improved internal resistance and cycle life," *JP:CS*, vol. 660, no. 1, p. 012009, 2015.

[53] M. Cowell, R. Winslow, Q. Zhang, J. Ju, J. Evans, and P. Wright, "Composite carbon-based ionic liquid supercapacitor for high-current micro devices," *JP:CS*, vol. 557, pp. 012061–5, Dec. 2014.

[54] A. Chen, D. Madan, P. K. Wright, and J. W. Evans, "Dispenser-printed planar thick-film thermoelectric energy generators," *J. of Micromechanics and Microengineering*, vol. 21, pp. 104006–9, Sept. 2011.

[55] B. P. Lechêne, M. Cowell, A. Pierre, J. W. Evans, P. K. Wright, and A. C. Arias, "Organic solar cells and fully printed super-capacitors optimized for indoor light energy harvesting," *Nano Energy*, vol. 26, pp. 631–640, Aug. 2016.

[56] D. Lee, G. Dulai, and V. Karanassios, " Survey of energy harvesting and energy scavenging approaches for on-site powering of wireless sensor- and microinstrument-networks," *SPIE Defense*, pp. 87280S–87280S–9, May 2013.

[57] V. Srinivasan and J. Weidner, "Mathematical Modeling of Electrochemical Capacitors," *Journal of The Electrochemical Society*, vol. 146, pp. 1650–1658, Nov. 1999.

[58] Y. V. Makarov, P. Du, M. C. W. Kintner-Meyer, C. Jin, and H. F. Illian, "Sizing energy storage to accommodate high penetration of variable energy resources," *IEEE Transactions on Sustainable Energy*, vol. 3, no. 1, pp. 34–40, 2012.

[59] R. Sioshansi, P. Denholm, T. Jenkin, and J. Weiss, "Estimating the value of electricity storage in PJM: Arbitrage and some welfare effects," *Energy Economics*, vol. 31, pp. 269–277, mar 2009.

[60] H. Mohsenian-Rad, "Coordinated Price-Maker Operation of Large Energy Storage Units in Nodal Energy Markets," *IEEE Transactions on Power Systems*, 2015.

[61] Department of Energy, "DOE Global Energy Storage Database."

[62] M. Korpaas, A. T. Holen, and R. Hildrum, "Operation and sizing of energy storage for wind power plants in a market system," *International Journal of Electrical Power & Energy Systems*, vol. 25, pp. 599–606, oct 2003.

[63] E. D. Castronuovo and J. a. P. Lopes, "Optimal operation and hydro storage sizing of a wind-hydro power plant," *International Journal of Electrical Power and Energy System*, vol. 26, no. 10, pp. 771–778, 2004.

[64] T. K. a. Brekken, A. Yokochi, A. V. Jouanne, Z. Z. Yen, H. M. Hapke, and D. a. Halamay, "Optimal energy storage sizing and control for wind power applications," *IEEE Trans. Sustain. Energy*, vol. 2, no. 1, pp. 69–77, 2011.

[65] Y. Ru, J. Kleissl, and S. Martinez, "Storage Size Determination for Grid-Connected Photovoltaic Systems," *IEEE Transactions on Sustainable Energy*, pp. 1–14, 2012.

[66] S. H. Madaeni, R. Sioshansi, and P. Denholm, "How thermal energy storage enhances the economic viability of concentrating solar power," *Proceedings of the IEEE*, vol. 10, pp. 335–347, 2012.

[67] E. Drury, P. Denholm, and R. Sioshansi, "The value of compressed air energy storage in energy and reserve markets," *Energy*, vol. 36, no. 8, pp. 4945–4973, 2011.

[68] X. Xi, R. Sioshansi, and V. Marano, "A stochastic dynamic programming model for co-optimization of distributed energy storage," *Energy Systems*, vol. 5, pp. 475–505, 2013.

[69] California Independent System Operator, "Open Access Same-time Information System (OASIS)."

[70] T. Thadewald and H. Büning, "Jarque-Bera test and its competitors for testing normality: A power comparison," *School of Business & Economics Discussion Paper: Economics*, vol. 2004/9, 2004.

[71] Energy Controls & Applications Laboratory, "Electricity Mapper."

[72] California ISO, "California Vehicle-Grid Integration (VGI) Roadmap," tech. rep., 2013.

[73] A. Hortaçsu and S. L. Puller, "Understanding strategic bidding in multi-unit auctions: A case study of the Texas electricity spot market," *RAND Journal of Economics*, vol. 39, no. 1, pp. 86–114, 2008.

[74] A. Kannan, U. V. Shanbhag, and H. M. Kim, "Addressing supply-side risk in uncertain power markets: stochastic Nash models, scalable algorithms and error analysis," *Optimization Methods and Software*, vol. 28, no. 5, pp. 1095–1138, 2013.

[75] A. Kannan, U. V. Shanbhag, and H. M. Kim, "Strategic behavior in power markets under uncertainty," *Energy Systems*, vol. 2, no. 2, pp. 115–141, 2011.

[76] D. Zhang, H. Xu, and Y. Wu, "A two stage stochastic equilibrium model for electricity markets with two way contracts," *Mathematical Methods of Operations Research*, vol. 71, no. 1, pp. 1–45, 2010.

[77] W. Tang and R. Jain, "Game-theoretic analysis of the nodal pricing mechanism for electricity markets," *Proceedings of the IEEE Conference on Decision and Control*, pp. 562–567, 2013.

[78] C. J. Day, B. F. Hobbs, and J. S. Pang, "Oligopolistic competition in power networks: A conjectured supply function approach," *IEEE Transactions on Power Systems*, vol. 17, no. 3, pp. 597–607, 2002.

[79] M. Ventosa, Á. Ì. Baíllo, A. Ramos, and M. Rivier, "Electricity market modeling trends," *Energy Policy*, vol. 33, no. 7, pp. 897–913, 2005.

[80] B. F. Hobbs, C. Metzler, and J. S. Pang, "Strategic gaming analysis for electric power networks: An MPEC approach," *IEEE Transactions on Power Systems*, vol. 15, no. 2, pp. 638–645, 2000.

[81] J. Barquín and M. Vázquez, "Cournot equilibrium in power networks," *Working Paper, Submitted to IEEE Transactions on Power Systems*, no. April, pp. 1–8, 2005.

[82] K. Neuhoff, J. Barquin, M. G. Boots, A. Ehrenmann, B. F. Hobbs, F. A. M. Rijkers, and M. Vazquez, "Network-constrained Cournot models of liberalized electricity markets: the devil is in the details," *Energy Economics*, vol. 27, no. 3, pp. 495–525, 2005.

[83] K. H. Lee, "Strategy equilibrium in Stackelberg model with transmission congestion in electricity market," *Journal of Electrical Engineering and Technology*, vol. 9, no. 1, pp. 90–97, 2014.

[84] J. Yao, S. S. Oren, and B. F. Hobbs, "Hybrid Bertrand-Cournot Models of Electricity Markets with Multiple Strategic Subnetworks and Common Knowledge Constraints," *Restructured Electric Power Systems: Analysis of Electricity Markets with Equilibrium Models*, pp. 167–192, 2010.

[85] O. L. Mangasarian and H. Stone, "Two-Person Nonzero-Sum Games and Quadratic Programming," *Journal of Mathematical Analysis and Applications*, vol. 9, no. 3, pp. 348–355, 1964.

[86] F. Wen and a. K. David, "Optimal bidding strategies for competitive generators and large consumers," *International Journal of Electrical Power and Energy System*, vol. 23, no. 1, pp. 37–43, 2001.

[87] J. Yao, S. S. Oren, and I. Adler, "Two-settlement electricity markets with price caps and Cournot generation firms," *European Journal of Operational Research*, vol. 181, no. 3, pp. 1279–1296, 2007.

[88] R. H. Kwon and D. Frances, "Optimization-Based Bidding in Day-Ahead Electricity Auction Markets: A Review of Models for Power Producers," in *Handbook of Networks in Power Systems I*, 2012.

[89] A. J. Conejo, M. Carrion, and J. M. Morales, *Decision Making Under Uncertainty in Electricity Markets*. New York, NY: Springer Science+Business Media, 2010.

[90] R. Rajagopal, E. Bitar, P. Varaiya, and F. Wu, "Risk-limiting dispatch for integrating renewable power," *International Journal of Electrical Power and Energy Systems*, vol. 44, no. 1, pp. 615–628, 2013.

[91] L. Wang, M. Mazumdar, M. D. Bailey, and J. Valenzuela, "Oligopoly models for market price of electricity under demand uncertainty and unit reliability," *European Journal of Operational Research*, vol. 181, no. 3, pp. 1309–1321, 2007.

[92] M. Dicorato, G. Forte, M. Trovato, and E. Caruso, "Risk-Constrained Profit Maximization in Day-Ahead Electricity Market," *IEEE Transactions on Power Systems on Power Systems*, vol. 24, no. 3, pp. 1107–1114, 2009.

[93] A. Baillo, M. Ventosa, M. Rivier, and A. Ramos, "Optimal offering strategies for generation companies operating in electricity spot markets," *IEEE Transactions on Power Systems*, vol. 19, no. 2, pp. 745–753, 2004.

[94] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, *Robust Optimization*. Princeton University Press, 2009.

[95] D. Bertsimas and M. Sim, "The Price of robustness," *Operations Research*, vol. 52, pp. 35–53, 2004.

[96] D. Bertsimas, D. B. D. Brown, and C. Caramanis, "Theory and Applications of Robust Optimization," *Operations Research*, p. 50, 2010.

[97] M. Aghassi and D. Bertsimas, "Robust game theory," *Mathematical Programming*, vol. 107, no. 1-2, pp. 231–273, 2006.

[98] R. Nishimura, S. Hayashi, and M. Fukushima, "Robust nash equilibria in N-person noncooperative games: Uniqueness and reformulation," *Pacific Journal of Optimization*, vol. 5, no. 2, pp. 237–259, 2009.

[99] A. Street, F. Oliveira, and J. M. Arroyo, "Contingency-Constrained Unit Commitment With Security Criterion: A Robust Optimization Approach," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1581–1590, 2011.

[100] R. Jiang, J. Wang, and Y. Guan, "Robust unit commitment with wind power and pumped storage hydro," *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 800–810, 2012.

[101] D. Bertsimas, E. Litvinov, X. A. Sun, J. Zhao, and T. Zheng, "Adaptive Robust Optimization for the Security Constrained Unit Commitment Problem," *IEEE Transactions on Power Systems*, vol. 28, pp. 1–8, 2013.

[102] T. Summers, J. Warrington, M. Morari, and J. Lygeros, "Stochastic optimal power flow based on convex approximations of chance constraints," *Proceedings - 2014 Power Systems Computation Conference, PSCC 2014*, vol. 72, pp. 116–125, 2014.

[103] L. Baringo and A. J. Conejo, "Offering strategy via robust optimization," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1418–1425, 2011.

[104] H. Haghighat, "Strategic offering under uncertainty in power markets," *International Journal of Electrical Power and Energy Systems*, vol. 63, pp. 1070–1077, 2014.

[105] C. Metzler, B. F. Hobbs, and J.-S. Pang, "Nash-Cournot Equilibria in Power Markets on a Linearized DC Network with Arbitrage: Formulations and Properties," *Networks and Spatial Economics*, vol. 3, pp. 123–150, 2003.

[106] B. F. Hobbs, E. Bartholomew, Y. Chen, G. Drayton, and W. Lise, "Improved transmission representations in oligopolistic market models," *2006 IEEE PES Power Systems Conference and Exposition*, pp. 81–86, 2006.

[107] J. Yao, I. Adler, and S. S. Oren, "Modeling and Computing Two-Settlement Oligopolistic Equilibrium in a Congested Electricity Network," *Operations Research*, vol. 56, no. 1, pp. 34–47, 2008.

[108] Y. Xie and U. V. Shanbhag, "On robust solutions to uncertain monotone linear complementarity problems (LCPs) and their variants," *53rd IEEE Conference on Decision and Control*, pp. 2834–2839, 2014.

[109] Y. Xie and U. V. Shanbhag, "On robust solutions to uncertain monotone linear complementarity problems and their variants," *Siam Journal of Optimization*, 2017.

[110] B. F. Hobbs, "Linear complementarity models of nash-Cournot competition in bilateral and POOLCO power markets," *IEEE Transactions on Power Systems*, vol. 16, no. 2, pp. 194–202, 2001.

[111] J. Slay and M. Miller, "Lessons Learned from the Maroochy Water Breach," in *Critical Infrastructure Protection* (E. Goetz and S. Shenoi, eds.), (Boston, MA), pp. 73–82, Springer US, 2008.

[112] R. Langner, "Stuxnet: Dissecting a cyberwarfare weapon," *IEEE Security and Privacy*, vol. 9, no. 3, pp. 49–51, 2011.

[113] A. Greenberg, "Unprecendented Malware Targets Industrial Safety Systems in the Middle East," dec 2017.

[114] A. Greenberg, "How an Entire Nation Became Russia's Test Lab for Cyberwar," jun 2017.

[115] R. M. Lee, M. J. Assante, and T. Conway, "Analysis of the cyber attack on the Ukrainian power grid," *SANS Industrial Control Systems*, p. 23, 2016.

[116] N. Perlroth and D. Sanger, "Cyberattacks Put Russian Fingers on the Switch at Power Plants, U.S. Says," mar 2018.

[117] K. Atherton, "It's not just elections: Russia hacked the US electric grid," mar 2018.

[118] B. Obama, "Executive Order 13636- Improving Critical Infrastructure Cybersecurity," 2013.

[119] B. Obama, "Presidential Policy Directive PPD-21 – Critical Infrastructure Security and Resilience," feb 2013.

[120] ISC, "Presidential Policy Directive 21 Implementation: An Interagency Security Committee White Paper," Tech. Rep. February, 2015.

[121] I. Rouf, H. Mustafa, M. Xu, and W. Xu, "Neighborhood watch: Security and privacy analysis of automatic meter reading systems," ... *Communications Security*, pp. 462–473, 2012.

[122] S. Sundaram and B. Gharesifard, "Distributed Optimization Under Adversarial Nodes," pp. 1–13, 2016.

[123] P. Blanchard, E. Mahdi El Mhamdi, R. Guerraoui, and J. Stainer, "Machine Learning with Adversaries: Byzantine Tolerant Gradient Descent," *Nips2017*, no. Nips, 2017.

[124] J. Zhang, P. Jaipuria, A. Chakrabortty, and A. Hussain, "A Distributed Optimization Algorithm for Attack-Resilient Wide-Area Monitoring of Power Systems: Theoretical and Experimental Methods," in *Decision and Game Theory for Security. GameSec 2014. Lecture Notes in Computer Science, vol 8840* (R. Poovendran and W. Saad, eds.), pp. 350–359, Springer, 2014.

[125] M. Liao and A. Chakrabortty, "Optimization Algorithms for Catching Data Manipulators in Power System Estimation Loops," *IEEE Transactions on Control Systems Technology*, pp. 1–16, 2018.

[126] C. Xie, O. Koyejo, and I. Gupta, "Generalized Byzantine-tolerant SGD," vol. 1, 2018.

[127] D. Yin, Y. Chen, K. Ramchandran, and P. Bartlett, "Byzantine-Robust Distributed Learning: Towards Optimal Statistical Rates," 2018.

[128] L. Su and N. Vaidya, "Fault-Tolerant Multi-Agent Optimization : Optimal Distributed Algorithms," in *Proceedings of the 2016 ACM Symposium on Principles of Distributed Computing*, vol. 1, (Chicago, Illinois), pp. 425–434, 2016.

[129] Y. Chen, S. Kar, and J. M. Moura, "Resilient Distributed Estimation Through Adversary Detection," *IEEE Transactions on Signal Processing*, vol. 66, no. 9, pp. 2455–2469, 2018.

[130] D. Alistarh, Z. Allen-Zhu, and J. Li, "Byzantine Stochastic Gradient Descent," pp. 1–20, 2018.

[131] S. Nabavi and A. Chakrabortty, "An Intrusion-Resilient Distributed Optimization Algorithm for Modal Estimation in Power Systems," *Conference on Decision and Control (CDC)*, no. Cdc, 2015.

[132] M. Liao and A. Chakrabortty, "A Round-Robin ADMM algorithm for identifying data-manipulators in power system estimation," *Proceedings of the American Control Conference*, vol. 2016-July, pp. 3539–3544, 2016.

[133] Y. Chen, S. Kar, and J. M. F. Moura, "Resilient Distributed Estimation: Sensor Attacks," pp. 1–8, 2017.

[134] G. K. Weldehawaryat, P. L. Ambassa, A. M. Marufu, S. D. Wolthusen, and A. V. Kayem, "Decentralised scheduling of power consumption in micro-grids: Optimisation and security," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10166 LNCS, pp. 69–86, 2017.

[135] S. Sundaram and B. Gharesifard, "Consensus-based distributed optimization with malicious nodes," *2015 53rd Annual Allerton Conference on Communication, Control, and Computing, Allerton 2015*, pp. 244–249, 2016.

[136] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Transactions on Information and System Security*, vol. 14, no. 1, pp. 1–33, 2011.

[137] X. Li, X. Liang, R. Lu, X. Shen, X. Lin, and H. Zhu, "Securing smart grid: Cyber attacks, countermeasures, and challenges," *IEEE Communications Magazine*, vol. 50, no. 8, pp. 38–45, 2012.

[138] M. Ozay, I. Esnaola, F. T. Y. Vural, S. R. Kulkarni, and H. V. Poor, "Sparse attack construction and state estimation in the smart grid: Centralized and distributed models," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 7, pp. 1306–1318, 2013.

[139] G. Liang, J. Zhao, F. Luo, S. R. Weller, and Z. Y. Dong, "A Review of False Data Injection Attacks Against Modern Power Systems," *IEEE Transactions on Smart Grid*, vol. 8, no. 4, pp. 1630–1638, 2017.

[140] H. Jan Liu, M. Backes, R. Macwan, and A. Valdes, "Coordination of DERs in Microgrids with Cybersecure Resilient Decentralized Secondary Frequency Control," vol. 9, pp. 2670–2679, 2018.

[141] V. Kekatos and G. B. Giannakis, "Distributed robust power system state estimation," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1617–1626, 2013.

[142] O. Vukovic and G. Dan, "Security of fully distributed power system state estimation: Detection and mitigation of data integrity attacks," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 7, pp. 1500–1508, 2014.

[143] C. Le Floch, F. Belletti, S. Saxena, A. M. Bayen, and S. Moura, "Distributed optimal charging of electric vehicles for demand response and load shaping," *Proceedings of the IEEE Conference on Decision and Control*, vol. 2016-Febru, no. Cdc, pp. 6570–6576, 2016.

[144] J. F. Mota, J. M. Xavier, P. M. Aguiar, and M. Püschel, "D-ADMM: A distributed algorithm for compressed sensing and other separable optimization problems," *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, no. 2, pp. 2869–2872, 2012.

[145] P. Sulc, S. Backhaus, and M. Chertkov, "Optimal Distributed Control of Reactive Power Via the Alternating Direction Method of Multipliers," *IEEE Transactions on Energy Conversion*, vol. 29, no. 4, pp. 968–977, 2014.

[146] S. C. Tsai, Y. H. Tseng, and T. H. Chang, "Communication-efficient distributed demand response: A randomized ADMM approach," *IEEE Transactions on Smart Grid*, vol. PP, no. 99, pp. 1–11, 2015.

[147] Y. Wang, L. Wu, and S. Wang, "A Fully-Decentralized Consensus-Based ADMM Approach for DC-OPF With Demand Response," *IEEE Transactions on Smart Grid*, pp. 1–11, 2016.

[148] New York State Energy Planning Board, "The Energy to Lead: New York State Energy Plan," tech. rep., New York State Energy Planning Board, 2015.

[149] North Carolina Clean Energy Technology Center, "Net Metering Policy Map," *DSIRE*, no. July, 2016.

[150] F. Flores-espino, "Compensation for Distributed Solar : A Survey of Options to Preserve Stakeholder Value," Tech. Rep. NREL/TP-6A20-62371, National Renewable Energy Laboratory, 2015.

[151] G. Meyers, "Black Hole Forms On Solar Net Metering In Nevada," jan 2016.

[152] Council GridWise Architecture, "GridWise Transactive Energy Framework Version 1.0," pp. 1–43, 2013.

[153] S. Chen and C.-c. Liu, "From Demand Response to Transactive Energy: the State-of-the-Art," *J. Mod. Power Syst. Clean Energy*, 2016.

[154] S. Borenstein, J. Bushnell, and F. Wolak, "Measuring Market Inefficiencies in California ' s Restructured Wholesale Electricity Market," *American Economic Review*, vol. 92.5, pp. 1376–1405, 2002.

[155] F. Wolak, "Lessons from the California Electricity Crisis," in *Electricity deregulation: choices and challenges*, ch. 4, 2005.

[156] California Public Utilities Commission, "Microgrids : A Regulatory Perspective," tech. rep., 2014.

[157] E. Munsing and S. J. Moura, "Security in Fully-decentralized Optimization Models," *In preparation*, 2018.

[158] A. Killeen, *Handbook of Digital Currency*. 2015.

[159] Deutsche Energie-Agentur GmbH (dena), "Blockchain in the energy transition: A survey among decision-makers in the German energy industry," tech. rep., 2016.

[160] J. Bonneau, A. Miller, J. Clark, A. Narayanan, J. A. Kroll, and E. W. Felten, "SoK: Research perspectives and challenges for bitcoin and cryptocurrencies," *Proceedings - IEEE Symposium on Security and Privacy*, vol. 2015-July, pp. 104–121, 2015.

[161] G. Wood, "Ethereum: a secure decentralised generalised transaction ledger," *Ethereum Project Yellow Paper*, pp. 1–32, 2014.

[162] D. T. Ton and M. A. Smith, "The U.S. Department of Energy's Microgrid Initiative," *The electricity Journal*, vol. 25, no. 8, pp. 84–94, 2012.

[163] A. Ahmad Khan, M. Naeem, M. Iqbal, S. Qaisar, and A. Anpalagan, "A compendium of optimization objectives, constraints, tools and algorithms for energy management in microgrids," *Renewable and Sustainable Energy Reviews*, vol. 58, pp. 1664–1683, 2016.

[164] L. I. Minchala-Avila, L. E. Garza-Castañón, A. Vargas-Martínez, and Y. Zhang, "A review of optimal control techniques applied to the energy management and control of microgrids," *Procedia Computer Science*, vol. 52, no. 1, pp. 780–787, 2015.

[165] C. Le Floch, F. Belletti, and S. Moura, "Optimal Charging of Electric Vehicles for Load Shaping : a Dual Splitting Framework with Explicit Convergence Bounds," *IEEE Transactions on Transportation Electrification*, vol. 7782, pp. 1–9, 2016.

[166] O. Sondermeijer, R. Dobbe, D. Arnold, and C. Tomlin, "Regression-based Inverter Control for Decentralized Optimal Power Flow and Voltage Regulation," in *Power & Energy Society General Meeting. IEEE*, 2016.

[167] S. Mhanna, A. C. Chapman, and G. Verbic, "A Fast Distributed Algorithm for Large-Scale Demand Response Aggregation," *IEEE Transactions on Smart Grid*, vol. 7, no. 4, pp. 2094–2107, 2016.

[168] E. Dall'Anese, H. Zhu, and G. B. Giannakis, "Distributed Optimal Power Flow for Smart Microgrids," *IEEE Transactions on Smart Grid*, vol. 4, no. 3, pp. 1464–1475, 2013.

[169] Q. Peng and S. H. Low, "Distributed Algorithm for Optimal Power Flow on an Unbalanced Radial Network," *53rd IEEE Conference on Decision and Control*, no. Cdc, pp. 0–7, 2015.

[170] S. Borenstein and J. Bushnell, "The U.S. Electricity Industry after 20 Years of Restructuring," *Annual Review of Economics*, vol. 7.1, pp. 437–463, 2015.

[171] J. D. Hertog, "Review of economic theories of regulation," *Tjalling C. Koopmans Institute Discussion Paper Series*, vol. 10, no. 18, pp. 1–59, 2010.

[172] M. Conoscenti, A. Vetro, and J. C. De Martin, "Blockchain for the Internet of Things: a Systematic Literature Review," no. November, 2016.

[173] J. Mattila, T. Seppala, C. Naucler, R. Stahl, M. Tikkanen, A. Badenlid, and J. Seppala, "Industrial Blockchain Platforms : An Exercise in Use Case Development in the Energy Industry," tech. rep., ETLA, 2016.

[174] M. Mihaylov, I. Razo-Zapata, R. RÇŐdulescu, S. Jurado, N. Avellana, and A. Nowé, "Smart grid demonstration platform for renewable energy exchange," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9662, pp. 277–280, 2016.

[175] L. P. Johnson, A. Isam, N. Gogerty, and J. Zitoli, "Connecting the Blockchain to the Sun to Save the Planet," dec 2015.

[176] Microgrid Media, ""It's Like the Early Days of the Internet," Blockchain-based Microgrid Tests P2P Energy Trading in Brooklyn," mar 2016.

[177] N. Li, "A Market Mechanism for Electric Distribution Networks," *Conference on Decision and Control (CDC)*, no. Cdc, pp. 2276–2282, 2015.

[178] M. E. Baran and F. F. Wu, "Optimal capacitor placement on radial distribution systems.," *IEEE Transactions on Power Delivery*, vol. 4, no. 1, pp. 725–734, 1989.

[179] M. Farivar and S. Low, "Branch Flow Model: Relaxations and Convexification, Part I," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2554–2564, 2013.

[180] E. M. Burger and S. J. Moura, "Generation following with thermostatically controlled loads via alternating direction method of multipliers sharing algorithm," *Electric Power Systems Research*, vol. 146, pp. 141–160, 2017.

[181] P. Scott and S. Thiébaux, "Dynamic Optimal Power Flow in Microgrids using the Alternating Direction Method of Multipliers," *CoRR*, pp. 1–8, 2014.

[182] Q. Peng and S. Low, "Optimal Branch Exchange for Distribution System Reconfiguration," *Arxiv Preprint*, p. 12, 2013.

[183] Ethereum Foundation, "Ethereum Frontier Guide," 2016.

[184] S. Diamond and S. Boyd, "{CVXPY}: A {P}ython-Embedded Modeling Language for Convex Optimization," *Journal of Machine Learning Research*, vol. 17, no. 83, pp. 1–5, 2016.

[185] A. Kosba, A. Miller, E. Shi, Z. Wen, and C. Papamanthou, "Hawk: The Blockchain Model of Cryptography and Privacy-Preserving Smart Contracts," in *Proceedings - 2016 IEEE Symposium on Security and Privacy, SP 2016*, 2016.

[186] S. Barber, X. Boyen, E. Shi, and E. Uzun, "Bitter to Better: How to Make Bitcoin a Better Currency,"

Appendix A

Decentralized Security- Analytic Solution

A.1 Analytic Solution Notes

The x-update step is an unconstrained quadratic program, and can be solved analytically. For clarity, we let $\gamma = Bz^k + u^k$ and proceed as:

$$\begin{aligned}
x^{k+1} &= \text{argmin}_x \quad x^T Px + c^T x + \frac{\rho}{2}||Ax + \gamma||_2^2 \\
x^{k+1} &= \text{argmin}_x \quad x^T Px + c^T x + \frac{\rho}{2}(Ax + \gamma)^T(Ax + \gamma) \\
x^{k+1} &= \text{argmin}_x \quad x^T Px + c^T x + \frac{\rho}{2}(x^T A^T Ax + 2\gamma^T Ax + \gamma^T\gamma) \\
0 &= \frac{\partial}{\partial x}\left(x^T Px + c^T x + \frac{\rho}{2}(x^T A^T Ax + 2\gamma^T Ax + \gamma^T\gamma)\right) \\
0 &= 2Px + c + \rho A^T Ax + \rho\gamma^T A \\
0 &= (2P + \rho A^T A)x + c + \rho\gamma^T A \\
x &= (2P + \rho A^T A)^{-1}(-c - \rho\gamma^T A)
\end{aligned}$$

Similarly, the z-update step can be solved analytically by letting $\mu = Ax^{k+1} - c + u^k$ and following a similar process to find:

$$z^{k+1} = (2Q + \rho B^T B)^{-1}(-d - \rho\mu^T B)$$

These analytic solutions are used in our implementation to avoid inaccuracies induced from a numeric solution.

Comparison with Central Solution

Because the problem is an unconstrained QP and entries with consensus between a subset of variables, a centralized solution can be computed by composing the cost matrices into a

single quadratic problem which can be solved analytically. This is shown here for the case where $c = 0$ and A and B are composed as described above, but also can be computed for other A, B.

We break P and Q into sub-matrices dependent on the number of consensus constraints p, where $P_{11}, Q_{00} \in \mathbb{R}^{p \times p}$ and the other dimensions follow accordingly.

$$P = \begin{bmatrix} P_{00} & P_{01} \\ P_{10} & P_{11} \end{bmatrix}$$
$$Q = \begin{bmatrix} Q_{00} & Q_{01} \\ Q_{10} & Q_{11} \end{bmatrix}$$
$$\Pi = \begin{bmatrix} P_{00} & P_{01} & 0 \\ P_{10} & P_{11} + Q_{00} & Q_{10} \\ 0 & Q_{10} & Q_{11} \end{bmatrix}$$

Similarly, the c and d vectors can be combined as

$$\kappa = \begin{bmatrix} c_0 \\ c_1 + d_0 \\ d_1 \end{bmatrix}$$

The problem can then be expressed as an unconstrained minimization problem:

$$\min_w \quad w^T \Pi w + \kappa^T w$$

which is solved by $w^* = -\frac{1}{2}(\Pi^T)^{-1}\kappa$

This central solution was used only for verification.