A Scalable Architecture for the Interconnection of Microgrids

by
Allison Fero
B.A. Physics
University of California (2013)
Submitted to the Institute for Data, Systems, and Society in partial fulfillment of the requirements for the degree of Master of Science in Technology and Policy at the MASSACHUSETTS INSTITUTE OF TECHNOLOGY June 2017
© Massachusetts Institute of Technology 2017. All rights reserved.
A Scalable Architecture for the Interconnection of Microgrids

by

Allison Fero

Submitted to the Institute for Data, Systems, and Society
on May 12, 2017, in partial fulfillment of the
requirements for the degree of
Master of Science in Technology and Policy

Abstract

Electrification is a global challenge that is especially acute in India, where about one
fifth of the population has no access to electricity. Solar powered microgrid technology
is a viable central grid alternative in the electrification of India, especially in remote
areas where grid extension is cost prohibitive. However, the upfront costs of microgrid
development, coupled with inadequate financing, have led to the implementation of
small scale, stand alone systems. Thus, the costs of local generation and storage are
a substantial barrier to acquisition of the technology. Furthermore, the issues of un-
certainty, intermittency, and variability of renewable generation are daunting in small
microgrids due to lack of aggregation. In this work, a methodology is provided that
maximizes system-wide reliability through the design of a computationally scalable
communication and control architecture for the interconnection of microgrids. An
optimization based control system is proposed that finds optimal load scheduling and
energy sharing decisions subject to system dynamics, power balance constraints, and
congestion constraints, while maximizing network-wide reliability. The model is first
formulated as a centralized optimization problem, and the value of interconnection
is assessed using supply and demand data gathered in India. The model is then for-
mulated as a layered decomposition, in which local scheduling optimization occurs at
each microgrid, requiring only nearest neighbor communication to ensure feasibility
of the solutions. Finally, a methodology is proposed to generate distributed optimal
policies for a network of Linear Quadratic Regulators that are each making decisions
coupled by network flow constraints. The LQR solution is combined with network
flow dual decomposition to generate a fully decomposed algorithm for finding the
dynamic programming solution of the LQR subject to network flow constraints.

Thesis Supervisor: Munther Dahleh
Title: William A. Coolidge Professor, Electrical Engineering and Computer Science

Thesis Supervisor: Mardavij Roozbehani
Title: Senior Research Scientist
Laboratory for Information and Decision Systems
Acknowledgments

I am incredibly thankful for the experience I have had at MIT over the last two years, and the amount of support that I received on the way. First I must thank my advisors Professor Munther Dahleh and Dr. Mardavij Roozbehani. The growth that I experienced while I was at MIT would not have been possible without them, and their depth of knowledge never fails to surprise me. It was a privilege to develop my research with such talented advisors. I’d like to thank the other members of LIDS for such a warm and stimulating community, especially Dr. Daria Madjidin, whose patience allowed for many multi-hour discussions. I’d also like to thank the Tata Center for Technology and Design, as well as Professor Rajeev Ram and Dr. Reja Amataya, for providing the resources that allowed me to explore otherwise theoretical ideas in such an important and applied domain as electrification in India.

Finally, I must thank my friends and family for being my foundation throughout this experience. The Technology and Policy Program attracts an incredible cohort of people who have been indispensable to my success at MIT. Thank you to my friends all over the world who never missed a beat, and even visited me across the country. Thank you to my sisters Kelly and Shelby and my brother Mostyn for the consistent group text updates, check-ins, and phone calls. Thank you Kyle for being there when I needed support of any kind, and being selfless without hesitation over the last couple years. And thank you to my parents, Mike and Katy, for all of the encouragement that got me to MIT and the unconditional love and guidance that inspire me every day.
# Contents

1 Electrification in India ............................ 9
   1.1 Motivation .............................................. 9
   1.2 Political Landscape ................................. 10
      1.2.1 Energy Policies ................................. 10
      1.2.2 Implementation ................................. 11
   1.3 Microgrids in India ................................. 13
   1.4 Microgrid Control Overview and Prior Art ........ 15
      1.4.1 Demand Response ................................. 15
   1.5 Energy Optimization in Offgrid India ............ 18
      1.5.1 Interconnection as a Solution ................. 19

2 Modeling Interconnection ......................... 21
   2.1 Constrained LQR ..................................... 24
   2.2 Centralized Solution ............................... 26
   2.3 Discussion and Simulations ....................... 28
      2.3.1 Centralized Simulations ....................... 29
      2.3.2 Discussion ....................................... 38

3 Layered Optimization for Local Implementation ... 41
   3.1 Dual Decomposition ................................. 42
   3.2 Notes on the Lagrangian ......................... 46
      3.2.1 Existence of a Unique Solution ............... 46
      3.2.2 Strong Duality .................................. 48
3.3 Subgradient Method .............................................. 50
  3.3.1 Sub-optimality bounds ...................................... 51
3.4 Fast Gradient Method ........................................... 52
3.5 Dual Decomposition Example ................................. 56

4 Dual Decomposition of the LQR Problem Under Network Flow Constraints 59
  4.1 The Model ....................................................... 61
  4.2 Policy Generation ............................................. 64
  4.3 Price update and Receding Horizon Control ............... 67
  4.4 Performance .................................................. 69

5 Conclusions and Future Work 73
  5.1 Overview ....................................................... 73
  5.2 Value of Interconnection ..................................... 73
  5.3 Decomposability of the Model ............................... 74
  5.4 Extension to the Distributed LQR ......................... 75
  5.5 Future Work ................................................ 75
  5.6 Final Remarks ............................................... 77
Chapter 1

Electrification in India

1.1 Motivation

Electricity access is a global issue with an estimated 1.2 billion people, or 17% of the global population, lacking access to electricity. A large portion of those without electricity reside in India, where about 20% of the country’s population, about 240 million people, remain un-electrified [1]. Although India represents about one-sixth of the global population, it uses only 6% of global energy [1]. However, India’s demographics and economy are rapidly evolving, with over a decade of strong economic growth. Alongside this economic growth there is an increased demand for energy. Given wage increases and population growth, electricity demand in India is expected to increase 4.4% annually from 2012 to 2040, tripling its overall demand [2]. India has shown a commitment to supplying a meaningful portion of this energy demand with renewable alternative energies, especially solar power. In its 2015-2016 federal budget, India has pledged to deploy 175 gigawatts of renewable energy by 2022, 100 of which will come from solar energy.

Solar powered microgrids are being installed throughout India in order to meet the goals of electrification through renewable generation, but often these microgrids are installed as independent, standalone systems. Due to the intermittent nature of renewable power such systems require either a large storage system (and therefore are quite expensive) or suffer from a lack of reliability. We propose a solution to mitigate
adverse behavior and maximize reliability through the design of a computationally scalable communication and control architecture for the interconnection of microgrids. By allowing microgrids to share their flexibility and excess supply and demand across a larger system, such a solution would smooth supply-demand imbalances, lead to more efficient use of energy, and reduce capital costs.

1.2 Political Landscape

1.2.1 Energy Policies

Utilities in India reside in a complex interface between politically mandated obligations and commercial goals. To understand the position of the utilities and the ability to onboard distributed renewable energy, some political context is necessary. Prior to 2003, State Electricity Boards (SEBs) were completely responsible for power generation, transmission, and distribution in India. With the passage of the Electricity Act of 2003, India fundamentally altered the power sector by unbundling the SEBs and de-licensing generation [3]. However, unbundling is only an option, and as of 2010, 11 states had yet to unbundle. The Electricity Act also established the Central Electricity Regulatory Commission (CERC), and the State Electricity Regulatory Commissions (SERCs). By 2011 all 29 states including Delhi had created state electricity regulators [4]. This means that each state has a uniquely structured power sector that responds to statewide regulation, but is also accountable to mandates passed at the national level.

Under the Electricity Act, universal access to electricity was also prioritized across the nation. The Act states that "State Government and the Central Government shall jointly endeavour to provide access to electricity to all." In order to provide for electrification of areas with low ability to pay, a system of cross subsidies was put in place. Cross-subsidization involves charging a subset of consumers higher rates than the cost of generation in order to apply discounted rates to other consumers with less ability to pay [5]. Competing with the subsidized price of power
is a fundamental challenge to implementing new electrification projects in India.

However, the Government of India does not hope to achieve universal access by means of centralized, conventional fuel alone. The mission purported by the central Government of India is also sustainability focused, pushing an agenda of renewable power generation that is administered by the Ministry of New and Renewable Energy (MNRE). For example, India’s Twelfth Five Year Plan (FYP) is scheduled to end in March, 2017. Over this period, the target is to add 29.8 GW of renewables, making total capacity 55 GW by 2017.

The MNRE’s goal of advancing the development of renewable technologies is spearheaded by the Jawaharlal Nehru National (JNN) Solar Mission. The JNN Solar Mission is broken down into three phases. The first phase was completed in 2012-2013. The goal of the first phase was to add 1,000 MW of solar capacity by focusing on the "low-hanging" solar thermal options and promoting off grid development [6]. Phase 2 takes place between 2013-2017, and should bring an additional 6,000 MW online. The main strategy for the first 3,000 MW in Phase 2 is the enforcement of the distribution company’s renewable purchase obligations (RPOs), as defined by the SERC in 2003. RPOs require a certain percentage of total electricity to be produced through renewable resources; states with low renewable potential are able to buy renewable energy certificates in lieu of production [7]. The second 3,000 MW should be enabled by international financing and technology transfer. Finally, phase 3 is planned from 2017-2022, bringing an additional 20,000 MW online, based on information from the first two phases.

1.2.2 Implementation

The goals of the Ministry of Power are important and have brought financing for renewable energy projects into the country. However, at the same time the conventional power sector in India continues to suffer a host of technical issues, challenging the effort for universal electricity access. Rapid expansion has left the grid overextended and unreliable. There are a series of instances in which the grid is underperforming:
• One fifth of India’s population does not have access to electricity [1]

• 9% of electricity demand went unmet in 2012 [2]

• For those who do have access, the limited goal of providing 6 hours a day of electricity has gone unmet. For example, in 2008-2009, Bihar received between 1.3 to 6.3 hours a day of electricity [3]

• Electricity distribution also suffers from enormous, costly losses. In 2010-2011 India’s nationwide losses were 23.97% [3]

Part of the problem is that the policies in place do not reflect the challenges of the actual system. For example, in many cases grid extension is infeasible in practice. In their 2014 review, Aggarwal et al. found that the price consumers paid to distribution companies per kWh was roughly 3 Rs. on average. However, the cost of generating, transmission, and distribution to remote areas between 5 and 25 km from the nearest grid would be between Rs. 3.18/kWh and 231/kWh [3]. Similarly, the lowest bid for a solar project in Phase 1 of the JNN Solar Mission was much lower than expected at Rs 7.49/kWh [8], but still more costly than conventional generation which is priced in the range of Rs 2-3/kWh. The goal of universal access from the 2003 Electricity Act is hampered by the cross subsidy policies laid out in the same act. Moreover, Phase 2 and 3 of the JNN Solar Mission are relying heavily on the fulfillment and enforcement of RPOs at the state level. However, because of the grid expansion and tariff structures, many distribution companies are financially stressed. Given the ailing state of the distributors, it is not certain whether regulators will be willing to further burden them with strict enforcement.

In contrast, the Solar Mission had success fulfilling its goals in Phase 1. The total share of renewable energy in India increased from 7.8% in 2008 to 12.3% in 2013 [8]. However, in the first phase, the metrics allowed for its goals to be achieved with conventional technology. Solar was installed off site, or in central thermal generation units. Cheap thermal energy from coal was bundled with more expensive solar generation to mask the high cost of solar [9]. As the phases of the JNN Solar Mission
get more ambitious, solar energy is expected to come online in all manners possible, including through distributed generation via microgrids. The introduction of novel microgrid technology that allows microgrids to operate more cost effectively has the potential to contribute to the success of both the JNN Solar Mission and the goal of electricity access for all in India.

1.3 Microgrids in India

There is a large diversity in what is considered a ‘microgrid’ in different contexts. In India, a microgrid can range anywhere in capacity from 200-500 W through upwards of 100kW. Microgrids are defined by onsite generation which can be provided by any number of resources such as biomass, hydro, wind, solar, diesel, or some combination. Microgrids do not require a renewable generation source, but there is a large potential for renewable integration in India because India has around 900 GW of commercially exploitable renewable energy sources, about 750 GW of which are solar [10]. Already 1.1 million households rely on solar energy for lighting needs, and roughly 51% of distributed renewable energy providers use solar as the main source of power [10] [11]. Microgrids are alternatives to conventional grid extension. Some advantages of microgrids include low upfront costs as compared to grid extension, limited maintenance, and the ability to incorporate renewables. On site generation and distribution also means transmission line losses are minimized.

However, even despite the relatively low capital investment required to develop a microgrid as compared to grid extension, cost is still a substantial barrier to entry for small microgrid companies. A larger 10 kW system can cost upwards of $30,000 dollars [11]. Companies attempting to finance microgrids must secure upfront capital as well as long term financing. Often these long term financing products are subject to high interest rates because of high perceived risk of entrepreneurs without a long financial history [11]. Entrepreneurs also have to contend with uncertainty of grid extension into rural areas, and communities inability to consistently pay for energy [11, 12]. Subsidies are available to offset large upfront costs, such as the "Rajiv
Gandhi Grameen Vidyutikaran Yojana” (RGGVY) scheme which subsidizes 90% of capital costs towards qualified rural electrification projects. Subsidizing upfront costs alone can incentivize developers to build microgrids without investing in long term operations [3,12].

These challenges, combined with bureaucratic challenges to receiving subsidies, have culminated with many entrepreneurs choosing to bypass subsidies and long term financing altogether, in favor of a business plan that is more tractable. In an assessment of solar home systems in Uttar Pradesh, one survey based study found that the willingness to pay for access to the system was about 1,828 rupees per year, which is roughly equivalent to $2.40 per month, well under many microgrids operating costs [13]. In order to overcome concerns regarding financing and uncertain payback, many existing microgrid companies choose to make very small systems that only offer lighting and cellular telephone charging at a competitive price to kerosene consumption. These systems cost around $1,000 to implement and thus recovery time is as little as 2-3 years [11]. For example, the microgrid company Mera Gao offers its 240 W system for around $900 and serves energy for lighting and cell phone charging for about 30 households. Sizing a system to be small has allowed companies to achieve success in regions such as the one surveyed in Uttar Pradesh, where consumers are looking to replace kerosene power with comparably priced electricity services. In Section 2.3, microgrid interconnection will be assessed under these types of microgrid assumptions.
1.4 Microgrid Control Overview and Prior Art

Microgrid implementation provides an array of novel technical challenges, regardless of the political surroundings. In general, control systems for microgrids can fall into three categories: primary, secondary and tertiary [14,15]. Primary controllers are fast and independent, allowing distributed generators (DG) to operate autonomously, while secondary and tertiary controllers support operations [15]. This work is concerned with the tertiary level of microgrid control. The focus is on the operation of the power system via energy sharing and management, in order to maximize systemwide efficiency and reliability.

The overarching goal in any power system is to match the demand for electricity with the supply at all times. Conventionally this balance is done on the supply side; system operators estimate forecasted demand and commit a certain amount of generation capacity accordingly. Standalone microgrids in India challenge this conventional approach, because renewable energy generation such as solar is exogenous and cannot be controlled. However, there are two controllable features that an isolated microgrid operator can influence. First, in lieu of controlling generation assets, standalone microgrid operators can choose how and when to dispatch one’s battery storage resource. Second, operators can partake in demand-side management, or demand response, in order to balance supply and demand. This section reviews existing literature on demand side management techniques, and contextualizes it to the case in India.

1.4.1 Demand Response

Demand response is the provision of incentives through pricing or another mechanism to reduce or offset peak demand through load shedding or shifting. There are two main mechanisms to incentive behavior changes amongst customers: direct participation with energy prices, and deferrable load scheduling.
Demand Management Through Energy Markets

The first major approach to implement demand response is to have consumers participate directly with the energy market. This concept can be generalized into demand response through dynamic retail pricing of electricity, in which price signaling is used to prompt voluntary load shedding or fulfillment. By providing a market mechanism, consumers ultimately make the decision on the priority of each load relative to the price of fulfillment. Many different pricing policies have been discussed in the literature. Borenstein et. al. compares the dynamics of real time pricing (RTP), in which the price is updated constantly, of electricity against critical peak pricing (CPP) and time-of-use (TOU). Time-of-use pricing is the simplest model, in which electricity prices vary throughout the day on a predetermined schedule, has already been implemented in some areas, and has been demonstrated as a means to implement intelligent dispatching with storage (Pepermans 2005, Hopkins 2012). However, the evaluation by Borenstein et. al. found that real-time pricing is most effective [16].

Enabling technology for real-time pricing has been studied extensively. Pipattanasomporn et. al. discusses the design of a multi-agent system that would allow for real-time price response through communication between agents [17]. Tsikalakis, et al. discussed the concept of a central microgrid central controller (MGCC) [15]. The MGCC’s function is to optimize the microgrid’s operation through participation in a real-time electricity market. Such a market based strategy would take bids from distributed generation units at fixed time intervals, as well as bids for supply or demand curtailment from customers, then optimize operation according to open market prices [18]. However, Roozbehani et. al. points out that there are challenges that must be addressed in the application of real-time pricing. Customer access and response to the real prices of electricity interacts with the market, creating a closed-loop feedback system that may increase system volatility and decrease robustness [19,20]. Pricing mechanisms and the associated control systems must be designed to appropriately address the issue of robustness in the system.
Another control solution to supply-demand balancing is by introducing demand side flexibility through ‘deferrable’ loads. Deferrable loads are loads that do not need to be served immediately, but must be satisfied within a fixed period of time. Often this technique is used with price response as the mechanism for incentivizing flexibility, although it does not have to be. Examples of possibly deferrable loads include temperature control, refrigeration, agricultural pumping, and ventilation. Such loads would behave similarly to storage resources from the perspective of the system operator [21,22], and could therefore decrease the need for system backup. Scheduling deferrable loads can be optimized in response to direct load control [23], or can be in response to dynamic pricing policies as described above [24], [18], [21].

Many deferrable demand strategies attempt to offset supply side variability from renewable generation with demand side flexibility in the form of deferred demand fulfillment to flexible customers using price mechanisms [18,24]. Papavasiliou et. al. proposes that a central scheduler should receive information about energy requests and their respective deadlines, optimizing resource allocation. Should the scheduler be unable to fulfill demand by its deadlines, power could be purchased from an electricity spot market. [24]. Neely et. al. also attempts to mitigate supply variability concerns with demand flexibility, using a Lyapunov optimization to improve performance [25]. Materassi et al. considers the scheduling of aggregate deferrable demand that must be satisfied by a finite time horizon, and derive a hierarchical coordination mechanism to achieve feasible policies [18]. Madjidian et al. shows that aggregated deferrable loads can emulate certain types of storage given appropriate mixed-slack policies. [22]. These models outline the potential for using load deferment in lieu of storage or expensive generation ramp-up to mitigate the effects of supply side uncertainty and balance supply and demand.
1.5 Energy Optimization in Offgrid India

The benefits of demand-side optimization have been well studied in countries such as the United States, but in India are just beginning to be seriously considered. A preliminary assessment of a smart grid test pilot conducted by the Tata Power Delhi Distribution Limited (TPDDL) in Delhi underscored the substantial benefits demand response can bring to the power sector, for both utilities and consumers [26]. In this case, TPDDL relied on non-price based signals; instead, customers received an automated signal from the utility and automatically reduced or increased demand, effectively shedding and shifting demand. The study found:

- On the utility side, DR provides savings by reducing unscheduled interchange (UI) withdrawals, lower purchase costs on the wholesale day-ahead market, and the ability to avoid generation from high-cost marginal generators
- An estimated between 4 to 6 Rs can be saved per kWh depending on when the demand is offset

In general, much of the literature in the field has been done in the United States and other nations with fully developed centralized generation. Thus, the available literature has done little to assess the value of energy management on microgrid systems in rural and off-grid areas. Solar powered microgrids are being installed throughout India in order to meet the goals of electrification through renewable generation, but often these microgrids are installed are small and independent, which are not adequately addressed by existing research (see Section 1.3). Immediately problems arise with the existing status quo on energy management. The two techniques described above, price response and load deferment, both have failings in the context of off-grid India. First, one issue with promoting demand management through price response is the possibility of inequity. Relying on demand elasticity through pricing could price some users out of the market, especially when the system isolated and is resource constrained. Although not a technical issue, the notion of preserving fairness in energy systems is important when attempting to provide universal access to energy. Second, demand response as a mechanism alone may provide limited flexibility given low
power consumption of typical residential loads in India. Finally, having deferrable loads with a fixed deadline often assumes the existence of types of loads that are not readily available in rural India, such as heating, cooling, and refrigeration. Furthermore, knowing a load’s ‘deadline’ requires data collection that is infeasible under existing circumstances in microgrid implementation.

1.5.1 Interconnection as a Solution

The solution proposed here is the interconnection of microgrids with each other or with the central grid. By allowing microgrids to share their flexibility and excess supply and demand across a larger system, such a solution should lead to more efficient use of energy, and reduce capital costs. Interconnection would take advantage of variations between individual microgrids to enable supply and demand matching on an aggregate level across the network. The system would have control not only of local storage dispatch and demand scheduling, but also energy sharing across the network. Interconnected demand scheduling and power sharing combines concepts from both real time pricing and load deferment in a way that is applicable to the implementation of microgrids in India, while addressing the issues discussed above. Explicit deadline constrained load deferment is a challenge given the size and type of loads typical in Indian microgrids. Instead, the scheduler is given the ability to defer loads by backlogging the energy demand and holding it over into the next period. This is allowed, but is penalized. Thus, all load is considered somewhat “deferrable,” but the backlog of unmet demand is minimized. Plus, microgrid interconnection would improve the flexibility and reliability of each grid in the network, at limited cost. In one estimate of solar mini grid costs for a very small, 100W capacity system, the battery, panels, and LED lights were 67% of total system costs, while cables, switches, and other miscellaneous installation costs were only 15% [27]. By incorporating interconnection, the same levels of reliability can be obtained with lower costs spent on materials and storage.
Chapter 2

Modeling Interconnection

In order to quantify the benefits of interconnection, the microgrid network can be modeled as an optimization problem over a discrete, finite time interval. A known network of \( n \) microgrids that are interconnected by \( m \) transmission lines can be represented as a graph in which each microgrid is represented by a node and the transmission lines are represented by edges. Each transmission line connects exactly two microgrids. It is possible to model this system as a connected, directed graph \( G = (N, E) \) with known topology of \( n \) vertices and \( m \) edges.

As discussed in Chapter 1, the goal of the microgrid system operator is to minimize the net imbalance between demand forecasted and the demand served at each node at all times. In the case that demand is unable to be served at any given time, it is assumed that the demand remains in the system until it is served or until the day is over, at the end of a 24 hour period. This is represented mathematically as a backlog...
state. The cost function has two terms in it. Costs are associated with the backlog of unmet demand and transmitting energy between microgrids across the transmission lines. The cost of backlog is quadratic so that larger deviations are penalized more heavily.

We keep track of two states. The first state, \( x(t) \in \mathbb{R}^n \) is a vector recording the total imbalance between energy served and energy demanded at every node. The imbalanced is backlogged, with a negative balance implying that there is unserved demand from earlier time periods. We require state \( x \) to be less than or equal to zero at all times which ensures that demand is not over-served. The second state is a vector of battery states \( b(t) \in \mathbb{R}^n \) at each node at time \( t \), which tracks the amount of storage available at each node \( i \). Each battery is limited by its maximum capacity and a positivity constraint.

We assume that we have some forecasted exogenous supply and demand, where \( d_i(t) \) represents aggregate local demand at node \( i \) and \( s_i(t) \) is aggregate local supply at node \( i \). Supply and demand forecasts are assumed to be given as positive values. Our decisions at every time point \( t \) are \( v_i(t) \): how much demand to serve locally, as well as \( f_{ij}(t) \ \forall (i,j) \in E \): the amount of energy to be traded with other microgrids in the network through the links. The demand served is constrained to be positive so as not to allow for the battery to draw energy back from the consumers. The flow variables, or amount of energy exchanged with the network at each node, are subject to the capacity constraints of the transmission line and to network flow conservation. In the directed graph, negative flow implies that energy is flowing against the direction of the link (for example \( f_{23} \leq 0 \) implies that energy flows from node 3 to node 2 in Figure 2-1). Thus the model is formulated as follows:

\[
\min_{v_i, f_{ij}} \sum_{i \in N} \sum_{t=1}^{T} Q_i x_i^2(t) + \sum_{(i,j) \in E} \sum_{t=1}^{T-1} R_{ij} f_{ij}^2(t) \tag{2.1}
\]

subject to:

\[
x_i(t + 1) = x_i(t) + v_i(t) - d_i(t) \tag{2.2}
\]
\[ b_i(t + 1) = b_i(t) + \sum_{j \in I(i)} f_{ji}(t) - \sum_{j \in O(i)} f_{ij}(t) - v_i(t) + s_i(t) \quad (2.3) \]

\[ v_i(t) \geq 0 \quad (2.4) \]

\[ x_i(t) \leq 0 \quad (2.5) \]

\[ b_i(t) \geq 0 \quad (2.6) \]

\[ b_i(t) \leq b_{max} \quad (2.7) \]

\[ -c \leq f_{ij}(t) \leq c \quad (2.8) \]

Note that the battery state \( b_i(t) \) has no cost associated with it. The role of the battery is to provide flexibility in our decision making at each time point. While decision variables \( v_i(t) \) are not explicitly upper-bounded, the lower bound on battery state \( b_i \) serves to dynamically provide a maximum for the energy available at time \( t \).

We can describe our system in a more compact form by defining new variables \( \tilde{x}, u, \) and \( w \):

\[ \tilde{x}(t) = [x_1(t) \quad b_1(t) \quad ... \quad x_N(t) \quad b_N(t)]^\top \]

\[ u(t) = [v_1(t) \quad ... \quad v_N(t) \quad f_{ij}(t)]^\top \]

\[ w(t) = [d_1(t) \quad s_1(t) \quad ... \quad d_N(t) \quad s_N(t)]^\top \]

Our problem can now be written:

\[ \min_u \tilde{x}(T)^\top Q \tilde{x}(T) + \sum_{t=1}^{T-1} \left\{ \tilde{x}(t)^\top Q \tilde{x}(t) + u(t)^\top Ru(t) \right\} \quad (2.9) \]

subject to:

\[ \tilde{x}(t + 1) = A \tilde{x}(t) + B u(t) + w(t) \quad (2.10) \]

\[ D \tilde{x}(t) - \gamma \leq 0 \]

\[ u \in U \]
Define

\[
D_i = \begin{bmatrix} 1 & 0 \\ 0 & -1 \\ 0 & 1 \end{bmatrix}, 
Q = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \gamma = \begin{bmatrix} 0 \\ 0 \\ -b_{\text{max}} \end{bmatrix}
\]

then D is the block diagonal matrix with blocks \(D_i\) and \(\gamma = [\gamma_i]^T\) for \(i = 1...N\). \(A\) is block diagonal matrix of \([2 \times 2]\) identity vectors, \(R\) is a diagonal matrix with entry 1 for \(f_{ij}\) and 0 for \(v_i\), and \(\hat{B}\) is the matrix of decision variable coefficients.

### 2.1 Constrained LQR

The linear quadratic regulator is a well studied dynamic programming problem. It refers to the optimization of a linear dynamic system over a quadratic cost function \(J(x_t, u_t)\):

\[
J(x_t, u_t) = \min_u \mathbb{E}_{w_t}\{x(T)^TQx(T) + \sum_{t=0}^{T-1} x^T(t)Qx(t) + u^T(t)Ru(t)\}
\]

subject to: \(x(t + 1) = Ax(t) + Bu(t) + w(t)\)

where matrix \(Q\) is assumed to be positive semi-definite and \(R\) is positive definite, and the expected value of disturbances \(w\) is assumed to be zero. In the case of a linear quadratic system with no other constraints, it is possible to find an analytical solution even under uncertain conditions [28]. The optimal control \(\mu_t(x_t)\) is found to be linear in \(x\) with the form

\[
\mu_t(x_t) = L_t x_t
\]

where \(L_t\) is determined by solving the discrete time Ricatti equation

\[
L_t = -(B^T K_{t+1} B + R)^{-1} B^T K_{t+1} A
\]

with \(K\) defined as

\[
K_t = Q + A^T K_{t+1} A - A^T K_{t+1} B (R + B^T K_{t+1} B)^{-1} B^T K_{t+1} A
\]
and the optimal cost is given by:

\[ J_0(x_0) = x_0^\top K_0 x_0 + \sum_{t=0}^{T-1} \mathbb{E}\{w_t^\top K_{t+1} w_t\} \] [28]

Our problem formulation given in Equation 2.9 is similar to the well studied linear-quadratic regulator (LQR) problem, in that we are minimizing a quadratic cost function of a system whose dynamics are described by a linear time invariant system. However, the optimization problem presented in Equation 2.1 has key distinctions. First, the addition of even simple constraints destroys the structure of the Ricatti equation, and can quickly make the classical dynamic programming approach computationally intractable [29]. Second, matrix R in Equation 2.9 is positive semidefinite, not positive definite.

The constrained LQR (CLQR) problem has also been studied extensively for deterministic problems. Given the deterministic CLQR problem:

\[
J(x_t, u_t) = \min_u x(T)^\top Q x(T) + \sum_{t=0}^{T-1} x^\top(t)Q x(t) + u^\top(t) R u(t)
\]

subject to: \( x(t+1) = A x(t) + B u(t) + w(t) \)

\( u \in U \)

It is possible to reformulate the LQR as a quadratic programming problem parametrized by the initial conditions \( x_o \) [30]. This can be shown directly by defining matrices:

\[
x = \begin{bmatrix} \tilde{x}(1) \\ \vdots \\ \tilde{x}(T) \end{bmatrix}, u = \begin{bmatrix} u(1) \\ \vdots \\ u(T) \end{bmatrix}, w = \begin{bmatrix} w(1) \\ \vdots \\ w(T) \end{bmatrix}
\]

\[
F = \begin{bmatrix} A \\ \vdots \\ A^T \end{bmatrix}, H = \begin{bmatrix} B \\ AB & B \\ \vdots & \vdots & \vdots \\ A^{T-1} B & A^{T-2} B & \cdots & B \end{bmatrix},
\]
\[ \tilde{Q} = \begin{bmatrix} Q & \cdots \\ \vdots & \ddots \\ Q & \cdots \end{bmatrix}, \tilde{R} = \begin{bmatrix} R \\ \vdots \\ R \end{bmatrix} \]

Re-writing the LQR in terms of \( F, H, \tilde{Q}, \) and \( \tilde{R}, \) one can equivalently write the problem as:

\[
\min_u x_o^T F^T \tilde{Q} F x_o + 2 x_o^T F^T \tilde{Q} H u + u^T (\tilde{R} + H^T \tilde{Q} H) u
\]

subject to: \( u \in U \)

Therefore, the CLQR problem can also be formulated as a quadratic program over some constraint set, and solved explicitly given some initial value \( x_0 \) [31, 32]. Bemporad et al. extended this to show that for any initial value, the CLQR can be formulated as a multi-parametric quadratic program and solved for all \( x_0 \) within a polyhedral set of values, leading to a piece-wise affine optimal control policy [33].

### 2.2 Centralized Solution

Instead of taking the dynamic programming approach, the problem can be formulated as a Quadratic Program (QP). Given initial values \( x_0 \) and \( b_0 \) at every node and known supply and demand forecasts, backlog state \( x_i \) and battery state \( b_i, \) as defined by equations 2.2 and 2.3, can each be explicitly written as \( T \) static constraints. Define vectors \( x_i \) and \( b_i: \)

\[
x_i = \begin{bmatrix} x_i(1) \\ \vdots \\ x_i(T) \end{bmatrix}, b_i = \begin{bmatrix} b_i(1) \\ \vdots \\ b_i(T) \end{bmatrix}
\]
Also, define L as the T*T lower triangular matrix of ones:

\[
L = \begin{bmatrix}
1 & 1 & 1 & \cdots \\
0 & 0 & 0 & \cdots \\
\vdots & \vdots & \vdots & \ddots \\
0 & 0 & 0 & \cdots & 1
\end{bmatrix}
\]

Then

\[
x_i = \bar{1}x_0 + L(v_i - d_i) \quad (2.11)
\]

\[
b_i = \bar{1}b_0 + L(-v_i + s_i - \sum_{j \in O(i)} f_{ij} + \sum_{j \in I(i)} f_{ji}) \quad (2.12)
\]

where each \(v_i\) and \(f_{ij}\) are \([(T - 1)\times 1]\) vectors:

\[
v_i = [v_i(1) \quad v_i(2) \quad \cdots \quad v_i(T - 1)]^T
\]

\[
f_{ij} = [f_{ij}(1) \quad f_{ij}(2) \quad \cdots \quad f_{ij}(T - 1)]^T
\]

Therefore, the model is written as:

\[
\min_{v_i, f_{ij}} \sum_{i \in N} x_i^T Q_i x_i + \sum_{(i, j) \in E} f_{ij}^T R_{ij} f_{ij} \quad (2.13)
\]

subject to:

\[
x_i = \bar{1}x_0 + L(v_i - d_i) \quad (2.14)
\]

\[
b_i = \bar{1}b_0 + L(-v_i + s_i - \sum_{j \in O(i)} f_{ij} + \sum_{j \in I(i)} f_{ji}) \quad (2.15)
\]

\[
v_i \geq 0 \quad (2.16)
\]

\[
x_i \leq 0 \quad (2.17)
\]

\[
b_i \geq 0 \quad (2.18)
\]

\[
b_i \leq b_{i \text{max}} \quad (2.19)
\]
This problem formulation can be solved online with a QP solver for a finite period of time $T$.

### 2.3 Discussion and Simulations

The value of microgrid interconnection relies on the heterogeneity of the network. Realistically, neighboring microgrids in off-grid India will be faced with very similar supply and demand distributions. Thus, it is reasonable to assume that the supply and demand forecasts at each microgrid will be highly correlated, and therefore be drawn from the same distribution $s_i \in s_o$ and $d_i \in d_o$. However, one way in which to ensure that the network is not homogenous is to vary the amount of storage capacity at each microgrid.

For the simulations, solar data was retrieved from the National Renewable Energy Laboratory (NREL) from Ranchi, India from January through December. An aggregate demand load profile was generated from information was collected from 5 residential households which had access to a cell charger, two LED light bulb, and a fan [34]. To generate multiple forecasts to use at each node, a random offset was

![Figure 2-2: Supply and Demand Data](image)

(a) Annual Power Output of 250 W Solar Panel, Ranchi, India

(b) Aggregated Demand, 5 households

Figure 2-2: Supply and Demand Data
added to the original distribution by sampling from a scaled uniform distribution with zero mean \( r \in R(-.5,.5) \). Call the original supply and demand forecasts \( s_o \) and \( d_o \) respectively, then:

\[
s_i = s_o + r_i \cdot s_o
\]

\[
d_i = d_o + r_i \cdot d_o
\]

Typical offsets can be seen in Figure 2-3.

![Figure 2-3](image)

(a) Randomly Offset Supply  
(b) Randomly Offset Demand

2.3.1 Centralized Simulations

Simulations are used to assess the benefits of interconnection under varied circumstances. In each simulation, the optimization period is over 24 hours. The optimization is run each day, then at the end of each day backlog is forgiven but battery state is remembered. There is one case in which the model can become infeasible: if there is excess generation capacity and the battery is full, the model will become infeasible. To overcome this issue, a decision to "spill" energy can be added to the battery constraint:

\[
b_i(t + 1) = b_i(t) + \sum_{j \in I(i)} f_{ji}(t) - \sum_{j \in O(i)} f_{ij}(t) - v_i(t) + s_i(t) - spill_i(t)
\]
where spill is always positive. This would be avoided unless where entirely necessary by adding a significantly large penalty in the cost term. For the simulations, spill with linearly penalized with a coefficient of 100,000.

There are two reasons why the microgrid system may become strained. First, upon inspection of the data it is clear that there is a surge in power usage from Month 2 through Month 7 when the temperature is higher. The upsurge in power consumption is mostly attributed to the use of fan, which are assumed to turn on after reaching an ambient temperature of 32°Celsius in the load profile generation [34]. The system can be sized such that this power usage is taken into account, or it can be sized to only serve “critical” demand, which in this case refers to LED lighting and cell phone charging.

Often solar microgrids being implemented in India are sized to support small, critical loads (See Section 1.3). For example, microgrid companies such as Mera Gao would outfit a village of 25-30 households with a 240W capacity solar panel. A small microgrid like this would generally be supported with a 12V, 100-150Ah battery. Transmission line capacity is dependent on a series of factors that are cable dependent. Antoniou et al analyzes low voltage DC distribution networks, including an analysis of total power capacity of 3-core cables under various DC configurations. The lowest total power for a DC unipolar cable was 61 kW [35]. We conservatively set the maximum transmission between the microgrids at 10 kW. Let $Q$ be the $nxn$ and $R$ be the $mxm$ identity matrices.

For the simulations, the supply data is taken directly from the 250W solar panel output data. This is roughly equivalent to the size of microgrids currently being implemented in rural India. The demand data is multiplied by five to represent aggregate load of 25 households, which is the size of a typical small microgrid, then an offset is added at each microgrid, as described above. Aggregate supply and demand for the year are shown in Figure 2-4.
Network Effects and Battery Placement

First we are interested in studying the effect of the network on the optimization results. As mentioned above, a realistic way to introduce heterogeneity into the network is through battery placement strategies. Here, we evaluate the impact of the location of batteries with respect to the network topology to identify the impact of relative position of microgrids with and without storage capacity. We assess the impact using a linear network of 20 microgrids, shown in Figure 2-5a. First, storage is increased from 0 Whr capacity to 72,000 Whr capacity (equivalent to 20 12V, 100Ahr batteries), but the battery bank is always placed at the end of the network (Node 1 - red). Second, batteries are placed in the middle of the network (Node 10 - green). Finally, batteries are distributed throughout the network, with 1200 Whr at each node with a battery. Results are shown in Figure 2-5.

This is an example of how the heterogeneity of the nodes in relation to the topology of the network has consequence. By placing the batteries at the end of the network, not only does average performance decrease overall, but power is distributed
less equitably. Figure 2-5c shows the disparity in demand served nodally given equal sized capacity distributed either at the end or evenly throughout the network. Demand at the nodes further from the batteries will only be served until the costs of transmission outweigh the costs of backlog. This can be explored further with an extreme case.

Say there are only two nodes on the system. We examine the extreme case in which Node 1 has an abundance of supply \((s_1 >> d_1, b_{\text{max}} = \infty)\) and a very large storage bank, while Node 2 has constrained supply \((\sum_{t=1}^{T}(s_2(t) - d_2(t) < 0)\) and no storage capacity. Following from Equation 2.13, the cost function can be written as:

\[
\min_{f_{12}, v_1, v_2} x_1^T Q_1 x_1 + x_2^T Q_2 x_2^2 + f_{12}^T R_{12} f_{12}
\]

where

\[
x_i = \bar{I} x_0 + L(v_i - d_i)
\]

Plugging in from 2.21 for \(x_1\) and \(x_2\), and assuming \(x_0 = 0\) for both microgrids we have:

\[
\min_{f_{12}, v_1, v_2} (v_1 - d_1)^T L^T Q_1 L(v_1 - d_1) + (v_2 - d_2)^T L^T Q_2 L(v_2 - d_2) + f_{12}^T R_{12} f_{12}
\]
subject to:

\[ L(v_1 - d_1) \leq 0 \]  
\[ L(v_2 - d_2) \leq 0 \]  
\[ b_1 = \bar{b}_0 + L(-v_1 + s_1 - f_{12}) \]  
\[ 0 = -v_2 + s_2 + f_{12} \]  
\[ v_1 \geq 0 \]  
\[ v_2 \geq 0 \]  
\[ -c \leq f_{12} \leq c \]

For convenience, define new variable \( \tilde{v}_i = v_i - d_i \)

Under the assumptions that \( s_1 \gg d_1 \), and \( b_{\text{max}} = \infty \), the battery constraint on Node 1, Equation 2.24, is not active. Let us also assume that the congestion constraint is not active. Without the active battery constraint, the optimization over \( v_1 \) is separable, and can be found to optimally be \( d_1 \) for all time points. This leaves us with decisions \( f_{12} \) and demand to serve at Node 2 \( v_2 \).

\[
\min_{f_{12}, \tilde{v}_2} \tilde{v}_2^\top L^\top Q_2 L \tilde{v}_2 + f_{12}^\top R_{12} f_{12}
\]

subject to:

\[ L \tilde{v}_2 \leq 0 \]
\[ 0 = -\tilde{v}_2 + s_2 - d_2 + f_{12} \]
\[ -\tilde{v}_2 + d_2 \leq 0 \]

Replacing \( \tilde{v}_2 \) with the equality constraint \( \tilde{v}_2 = s_2 - d_2 + f_{12} \) and taking the gradient with respect to decision variables \( f_{12} \):

\[ 2L^\top Q_2 L(s_2 - d_2 + f_{12}) + 2R_{12} f_{12} \]
Applying KKT Conditions where $\mu \geq 0$:

$$-L^\top Q_2 L (s_2 - d_2 + f_{12}) - R_{12} f_{12} = \frac{1}{2} (\mu_1 L - \bar{\mu}_2)$$

$$f_{12} = -(L^\top Q_2 L + R_{12})^{-1} \left(\frac{1}{2} (\mu_1 L - \bar{\mu}_2) + L^\top Q_2 L (s_2 - d_2)\right)$$

In the extreme case that $s_2 - d_2$ is always negative, for example if $s_2 = 0$, then the negativity constraint is not active at the optimal solution, and we have:

$$f_{12} = (L^\top Q_2 L + R_{12})^{-1} (L^\top Q_2 L d_2)$$

The impact of the network is captured in the cost of transmitting energy between the microgrids. There is a trade-off when removing local resources in favor of a networked system; the overall system costs may be reduced, but individual microgrids whose utilities are locally defined only by unmet backlog could suffer. This stresses the importance of tuning the parameters $Q$ and $R$ to reflect the value of microgrids. In this example, in the case that backlog is weighted very heavily, i.e., $Q_2 >> R_{12}$, then all demand can be served. However, the larger the relative cost of transmission, the less demand will be served despite the generation capacity to do so.

**Performance As a Function of Batteries in Network**

Microgrid interconnection has the potential to lessen the need for on-site storage throughout the network. In this section, different storage capacities are compared for varying microgrid networks. Battery capacity is measured by total capacity in Whr, or by “battery saturation,” which we define by the number of microgrids that are outfitted with a 12V 100Ahr battery out of the total number of microgrids. Batteries are placed such that they have an equal number of nearest nodes without batteries at an equal distance. For the initial analysis, Month 7 is used as a representative month in which the system is strained (during the end of monsoon season) but without an excess of non-critical load.

To test the effects of battery saturation directly, six graphs are simulated with
various battery saturations (see Figure 2-7). There was little difference between the network configurations because batteries were dispersed throughout the network. In each situation, there were some clear trends. First, the value of interconnection first increases then decreases in number of batteries removed from the network. Second, and importantly, the overall reliability of the network does not immediately decrease. This implies that the storage capacity at each microgrid is not being fully utilized. Therefore, until a certain point, microgrid interconnection can replace physical batteries in the network without significantly retracting from reliability.

![Network Diagrams](image)

**Figure 2-6: Microgrid Networks Compared**

**Year-round Simulations**

As shown above, batteries that are evenly dispersed throughout a network will perform comparably. Thus, for the remaining analysis we use the 10-node linear network (Network 6). In this section, the same analysis is performed on the 10-node linear network over the course of the entire year. Simulating over the entire year includes more demand side variance because of the high levels of non-critical fan loads during hotter months. Figure 2-8 shows that even with high demand variance, a 40% decrease
in storage capacity only amounts to a 3% decrease in annual reliability. This is compared to losing only about 2% reliability for a 60% decrease in storage capacity during Month 7. Figure 2-9 compares the average backlog and battery states for all of the microgrids with and without interconnection. When all microgrids have batteries, there is very little difference between the backlog and battery state at any given time. However, Figure 2-9b shows the backlog state without interconnection often exceeds the backlog state with interconnection, especially outside of Months 2-6. The ability to charge storage capacity from excess capacity from neighboring microgrids in the interconnected system is also reflected by the battery differential in the top half of Figure 2-9b.

**Performance Under Uncertain Supply**

The optimal policy is found under deterministic assumptions on supply and demand. In some cases, the optimal policy may become suboptimal or infeasible depending on actual supply and demand. Take the case in which the amount of actual supply $\hat{s}_i(t)$ is different than predicted supply $s_i(t)$. In this case, the optimal policy must be modified to maintain feasibility. To ensure feasibility, the actual amount of demand
served locally $\hat{v}_i(t)$ can be adjusted at any given time:

$$\hat{v}_i(t) = \min(v_i(t), b_i(t - 1) + B_i u^*(t) + \hat{s}_i(t))$$
Then, there is a question as to whether the policy $u^*$ still performs better than the policy to greedily consume.

To test this, take Network 6 with 1200 Whr batteries at every other node (50% saturation). 30 days are selected randomly throughout the year, and the simulation is repeated 20 times on each day. On each day, the optimal policy is chosen given some known supply forecast $s$. Then the supply is given a uniformly distributed offset $r \in R(-.5,.5)$ centered around $s$ with magnitude of $alpha$

$$\hat{s} = s + \alpha (r * s)$$

where $\alpha$ ranges from 0 to 1 in increments of .1. Figure 2-10 shows the amount of demand served for varied levels of supply compared to the average demand served over the 30 days with no interconnection. Despite the fact that the policy is generated under deterministic assumptions, even with supply side uncertainty microgrid interconnection performs better than the independent system.

### 2.3.2 Discussion

The implications of this finding are that as currently implemented, stand-alone microgrids are not using the full capacity of the on-site storage. In the systems we
simulated, giving each system a 12V, 100 Ahr battery provides excess capacity. This storage capability could instead be replaced by interconnection with a network of existing microgrids. Whether or not a small microgrid should be outfitted with a battery would depend on the exact network configuration of the microgrid network, and which microgrids already had batteries. Overall, microgrid interconnection was shown to provide value even for microgrids that are near by and would have highly correlated supply and demand. During Month 7 when the demand was mostly critical load and supply was reduced due to monsoon season, it was possible to remove storage capacity from 6/10 of the microgrids with only a 2% reduction in reliability. This means that for microgrids that are sized to serve critical demand, just 40% of the microgrids could be outfitted with batteries without significantly detracting from reliability of critical load. Simulating over the entire year which included both critical and non-critical load, Figure 2-8 shows that even with high demand variance, a 40% decrease in storage capacity only amounts to a 3% decrease in annual reliability. Depending on the sizing and types of load on the system, microgrid networks could greatly reduce the total number of storage devices system-wide.
Chapter 3

Layered Optimization for Local Implementation

The centralized approach gives a sense of the value that microgrid interconnection can bring to a network of nearby villages. This problem can be directly solved via a centralized optimization as long as the controller has access to full information across all microgrids. However, it is often the case that microgrids would not have full access to each other’s information for privacy or communication reasons. Furthermore, it may be unrealistic to assume that a centralized controller would exist and that the other microgrids would be comfortable relinquishing decision making to a single controller. Because of this we are interested in formulating this problem in a decomposed way. The goal is to allow microgrids to retain their own proprietary supply and demand data, while still participating in the networked system. Instead of one global optimization, microgrids should make decisions autonomously, then participate in a marketplace to adjust the response to be globally optimal. To do this, the optimization problem is decomposed into a 2-layer scheme. The local “scheduling” layer encompasses on-site decision making on the amount of demand to serve locally, and the second, “congestion” layer ensures the feasibility of the local results by updating price signals.
3.1 Dual Decomposition

To compute the optimizations locally, we use a decomposition method known as dual decomposition. Dual decomposition is a layered optimization technique that relies on taking the Lagrangian dual of a subset of the constraint set. In general, the Lagrangian dual is formulated as follows. Given a problem

$$\min_u f_0(u)$$

subject to:

$$f_i(u) \leq 0 \quad i = 1, \ldots, m$$ (3.1)

$$h_j(u) \leq 0 \quad j = 1, \ldots, p$$ (3.2)

To formulate the Lagrangian $L: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ with constraint (3.1), introduce Lagrange multipliers $\lambda_i$:

$$L(u, \lambda) = f_0(u) + \sum_{i=1}^{m} \lambda_i^\top f_i(u)$$

Define the Lagrangian dual function $g: \mathbb{R}^m \to \mathbb{R}$

$$g(\lambda) = \min_u L(u, \lambda) = f_0(u) + \sum_{i=1}^{m} \lambda_i^\top f_i(u)$$

subject to: $h_j(u) \leq 0 \quad j = 1, \ldots, p$
The Lagrangian dual function gives us a lower bound on the optimal value of the original optimization problem [36]. In order to find the best lower bound one then solves the optimization problem

\[
\max_{\lambda} g(\lambda) \\
\text{subject to: } \lambda_i \geq 0 \quad i = 1, \ldots, m
\]

In dual decomposition, the cost function of the problem is separable in decision variables \( u \), but decisions are coupled by a subset of the constraint set. The Lagrangian dual is applied to the coupling constraints in order to decompose the problem into independent subproblems. A secondary, “master” problem then finds the optimal value for the dual vector. This can be interpreted as the setting of prices for the coupled resources that are optimized in each subproblem in order to guarantee feasibility of the primal problem [37].

Upon inspection, it is clear that our model may lend itself to a decomposition method because the cost function is fully separable in \( i \). However, the constraints are coupled by the flow variables \( f_{ij} \) and \( f_{ji} \) in Equation 2.3. In order to decompose this problem, we define a new variable \( k_{ji} \) where

\[
k_{ji} = f_{ji} \quad \forall j \in I(i)
\]  

(3.3)

The problem is now written as:

\[
\min_{v,f_{ij,k}} \sum_{i \in N} x_i^T Q_i x_i + \sum_{j \in O(i)} f_{ij}^T \hat{R}_{ij} f_{ij} + \sum_{j \in I(i)} k_{ji}^T \hat{R}_{ji} k_{ji}
\]  

(3.4)

subject to:

\[
x_i = \bar{1}x_0 + L(v_i - d_i)
\]  

(3.5)

\[
b_i = \bar{1}b_0 + L(-v_i + s_i + \sum_{j \in I(i)} k_{ji} - \sum_{j \in O(i)} f_{ij})
\]  

(3.6)

\[
v_i \geq 0
\]  

(3.7)
\[ x_i \leq 0 \]  
\[ b_i \geq 0 \]  
\[ b_i \leq b_{imax} \]  
\[ -c \leq f_{ij} \leq c \]  
\[ k_{ji} = f_{ji} \quad \forall j \in I(i) \]  

This problem is completely decomposable except for Equation 3.12. To eliminate the coupling constraint we formulate the Lagrangian dual problem:

\[ \mathbb{L}(\lambda, v, f, k) = \max_{\lambda} \min_{k_{ji}, f_{ij}, v_i} \sum_{i \in N} x_i^\top Q_i x_i + \sum_{j \in O(i)} f_{ij}^\top \hat{R}_{ij} f_{ij} + \sum_{j \in I(i)} \{ k_{ji}^\top \hat{R}_{ji} k_{ji} + \lambda_{ji}^\top (k_{ji} - f_{ji}) \} \]  
\[ \lambda_{ji} = [\lambda_{ji}(1) \quad \lambda_{ji}(2) \ldots \lambda_{ji}(T - 1)]^\top \quad \forall j \in I(i) \]  

Define \( g(\lambda) \) as

\[ g(\lambda) = \min_{k_{ji}, f_{ij}, v_i} \sum_{i \in N} [x_i^\top Q_i x_i + \sum_{j \in O(i)} f_{ij}^\top \hat{R}_{ij} f_{ij} + \sum_{j \in I(i)} \{ k_{ji}^\top \hat{R}_{ji} k_{ji} + \lambda_{ji}^\top (k_{ji} - f_{ji}) \}] \]  

We first compute \( g(\lambda) \) subject to constraints 3.5 through 3.11 to find \( f(\lambda) \) and \( k(\lambda) \).

The master problem is then:

\[ \max_{\lambda} g(\lambda) \]  

Equation 3.14 is not yet fully separable. The battery constraint, Equation 3.6, only contains flow variables \( f_{ij} \) for all \( j \in O(i) \), but the cost function is over all flow variables \( f_{ij} \) for all \((i, j) \in E\).

**Proposition 1** \( g(\lambda) \) can be fully decoupled across nodes, written as:

\[ g(\lambda) = \sum_{i \in N} \min_{f_{ij}, v_i} \left[ x_i^\top Q_i x_i + f_{ij}^\top \hat{R}_{ij} f_{ij} + k_{ji}^\top \hat{R}_{ji} k_{ji} + \lambda_{ji}^\top k_{ji} - \lambda_{ij}^\top f_{ij} \right] \]  

44
Proof: We can show this directly. First, 
\( g(\lambda) \) can be written as:

\[
g(\lambda) = \min_{k_{ji}, f_{ij}, v_i} \sum_{i \in N} [x_i^T Q_i x_i + \sum_{j \in O(i)} f_{ij}^T \hat{R}_{ij} f_{ij} + \sum_{j \in I(i)} \{k_{ji}^T \hat{R}_{ji} k_{ji} + \lambda_{ji}^T (k_{ji} - f_{ji})\}] = \sum_{i \in N} \min_{f_{ij} \forall j \in O(i), k_{ji} \forall j \in I(i)} \sum_{i \in N} \sum_{j \in I(i)} (\lambda_{ji}^T k_{ji} - \lambda_{ji}^T f_{ji})\]

(3.17)

We can show that the second term can also be written in terms of local decision variables \( f_{ij} \) for all \( j \in O(i) \) and \( k_{ji} \) for all \( j \in I(i) \):

\[
\min_{k_{ji}, f_{ij}} \sum_{i \in N} \sum_{j \in I(i)} (\lambda_{ji}^T k_{ji} - \lambda_{ji}^T f_{ij}) = \min_{k_{ji}, f_{ij}} \sum_{i \in N} \sum_{j \in I(i)} \lambda_{ji}^T k_{ji} - \sum_{i \in N} \sum_{j \in O(i)} \lambda_{ij}^T f_{ij} = \sum_{i \in N} \min_{f_{ij} \forall j \in O(i), k_{ji} \forall j \in I(i)} \lambda_{ji}^T k_{ji} - \lambda_{ij}^T f_{ij}
\]

(3.18)

Putting the two terms together, we have

\[
g(\lambda) = \sum_{i \in N} \min_{f_{ij} \forall j \in O(i), k_{ji} \forall j \in I(i)} \sum_{i \in N} \sum_{j \in I(i)} \sum_{i \in N} \sum_{j \in O(i)} \{x_i^T Q_i x_i + f_{ij}^T \hat{R}_{ij} f_{ij} + k_{ji}^T \hat{R}_{ji} k_{ji} + \lambda_{ji}^T (k_{ji} - f_{ji})\}
\]

(3.19)

This means that each node \( i \) optimizes its own total “inflow” \( k_{ji}(t) \) and chooses each “outflow” \( f_{ij}(t) \) based on price variables \( \lambda_{ij} \) and \( \lambda_{ji} \), which are communicated from neighboring microgrids. Note that “inflow” and “outflow” are defined by graph structure and do not represent power inflow and outflow - each can be either positive or negative.

**Proposition 2** Each entry of the matrix \( \hat{R}_{ij} \) is equal to one half of the corresponding entry of the matrix \( R_{ij} \).
Proof: The links of the graph $f_{ij}$ can be equivalently written as:

$$
\sum_{(i,j) \in E} f_{ij} R_{ij} f_{ij} = \sum_{i \in N} \sum_{j \in O(i)} f_{ij} R_{ij} f_{ij} = \sum_{i \in N} \sum_{j \in O(i)} f_{ij} R_{ij} f_{ij}
$$

$$
= \frac{1}{2} \sum_{i \in N} \sum_{j \in O(i)} f_{ij} R_{ij} f_{ij} + \frac{1}{2} \sum_{i \in N} \sum_{j \in I(i)} f_{ij} R_{ij} f_{ij}
$$

(3.20)

Then if $\hat{R}_{ij} = \frac{1}{2} R_{ij}$:

$$
\sum_{i \in N} \sum_{j \in O(i)} f_{ij} \hat{R}_{ij} f_{ij} + \sum_{j \in I(i)} k_{ji} \hat{R}_{ji} k_{ji} = \sum_{i \in N} \sum_{j \in O(i)} \frac{1}{2} f_{ij} R_{ij} f_{ij} + \sum_{j \in I(i)} k_{ji} \frac{1}{2} R_{ji} k_{ji}
$$

(3.21)

Since $k_{ji} = f_{ji}$ for all $j \in I(i)$, Equation 3.21 is equivalent to Equation 3.20.

3.2 Notes on the Lagrangian

3.2.1 Existence of a Unique Solution

A generic quadratic function $f(x) = 1/2 x^T P x + q^T x + r$ is strictly convex if and only if the matrix $P$ is positive definite ($P > 0$) [36]. We show here that both the dual and primal problem are strictly convex. Recall that the primal cost function is given by

$$
\min_{v_{i, f_{ij}}} \sum_{i \in N} x_i^T Q_i x_i + \sum_{(i,j) \in E} f_{ij} R_{ij} f_{ij}
$$

where $R_{ij} > 0$ and $Q_i > 0$. We can plug in $x_i$ defined by

$$
x_i = \tilde{1} x_0 + L(v_i - d_i)
$$

to formulate the cost function in terms of $x_0$:

$$
\min_{v_{i, f_{ij}}} \sum_{i \in N} T x_0^2 + (v_i - d_i)^T L^T L(v_i - d_i) + 2 \tilde{x}_0 L(v_i - d_i) + \sum_{(i,j) \in E} f_{ij} R_{ij} f_{ij}
$$

(3.22)
Redefine $\bar{v}_i = v_i - d_i$ and define decision vector

$$u(t) = [\bar{v}_1(t) \ ... \ \bar{v}_N(t) \ f_{ij}(t)]^\top$$

Then we have

$$f(u) = \min_u T x_0^2 + uHu + 2\bar{x}_0Fu$$

(3.23)

where $H$ as the positive definite block diagonal matrix with

$$
H = \begin{bmatrix}
L^\top L & & \\
& \ddots & \\
& & L^\top L \\
& & & R
\end{bmatrix}
$$

with $n$ $L^\top L$ blocks, and where $R$ is an $mxm$ diagonal matrix of $R_{ij}$. This implies the cost function $f(u)$ is strongly convex.

The decomposed problem is also strongly convex. In section 3, we decomposed the problem by adding new variables $k_{ji}(t) = f_{ji}(t)$ for all outgoing links $j \in I(i)$. As shown in Equation 3.20, $\bar{R} = \frac{1}{2}R$, so the cost function can be equivalently written as

$$
\sum_{i \in N \ v, f_{ij} \forall j \in O(i), k_{ji} \forall j \in I(i)} \min_{u} T x_0^2 + (v_i - d_i)^\top L^\top L(v_i - d_i) + \frac{1}{2} f_{ij}^\top R_{ij} f_{ij} + \frac{1}{2} k_{ji}^\top R_{ij} k_{ji} + 2\bar{x}_0 L(v_i - d_i)
$$

(3.24)

Define decision vector at each node:

$$u_i = [\bar{v}_i \ f_{ij} \ k_{ji}]^\top$$

Re-write the cost in terms of decision vectors $u_i$:

$$f_d(u) = \sum_{i \in N} \min_{u_i} T x_0^2 + u_i H_i u_i + 2\bar{x}_0 F_i u_i$$

(3.25)
where \( u \) is a vector of vectors \( u_i \) and \( H_i \) is the positive definite block diagonal matrix

\[
H_i = \begin{bmatrix}
L^\top L & \frac{1}{2} R_{ij} \\
\frac{1}{2} R_{ij} & \ddots \\
& \frac{1}{2} R_{jj}
\end{bmatrix}
\]

Therefore, the decomposed problem \( f_d(u) \) is strongly convex.

### 3.2.2 Strong Duality

We show strong duality between the primal, centralized optimization problem and the Lagrangian dual problem. In general, the Lagrangian dual may have some positive valued duality gap \( p^* - d^* \).

**Proposition 3** For our Lagrangian dual decomposition defined in Equation 3.13, the optimal duality gap between the primal optimal solution and the dual optimal solution is zero: \( d^* = p^* \)

**Proof:** Recall from equation 3.13 that the Lagrangian dual problem is a function of variables \( f, v, k \). We call the set of primal variables \( \mu \in U \) where \( U \) is defined by constraints 3.5 through 3.11.

From [36], we can express strong duality as the equality:

\[
\max_{\lambda} \min_{\mu \in U} L(\lambda, \mu) = \min_{\mu \in U} \max_{\lambda} L(\lambda, \mu) \tag{3.26}
\]

The left hand side of this equation is the master problem:

\[
\max_{\lambda} g(\lambda)
\]

Where \( g(\lambda) \) is defined by equation 3.14:

\[
g(\lambda) = \min_{k_{ji}, f_{ij}, v_i} \sum_{i \in N} [x_i^\top G_i x_i + \sum_{j \in O(i)} f_{ij}^\top R_{ij} f_{ij} + \sum_{j \in I(i)} \{k_{ji}^\top \hat{R}_{ji} k_{ji} + \lambda_{ji}(k_{ji} - f_{ji})\}]
\]
Because $\lambda$ is unconstrained, the only two solutions to this problem are

$$
\max_{\lambda} g(\lambda) = \begin{cases} 
  x^*_iTQi^*_i + f^*_i\hat{R}_{ij}f^*_i + k^*_j\hat{R}_{ji}k^*_j, & \text{if } k^*_j - f^*_j = 0 \quad \forall i \\
  \infty, & \text{otherwise}
\end{cases}
$$

Similarly, the right hand side of equation 3.26 is given by

$$
\min_{f_{ij}, k_{ij}, v_i} \sum_{i \in N} [x_i^T Q_i x_i + \sum_{j \in O(i)} f_{ij}^T \hat{R}_{ij} f_{ij} + \sum_{j \in I(i)} \{k_{ji}^T \hat{R}_{ji} k_{ji} + \lambda_{ji}^* (k_{ji} - f_{ji})\}] \quad (3.27)
$$

Once again, given arbitrary $\lambda^*_i$, equation (3.13) is only feasible if $k_{ji} - f_{ji} = 0$, and the optimal value is either $x_i^* T Q_i x_i + f_{ij}^* T \hat{R}_{ij} f_{ij} + k_{ji}^* T \hat{R}_{ji} k_{ji}^*$ or $-\infty$. Therefore,

$$
\max_{\lambda} \min_{\mu \in U} L(\lambda, \mu) = \min_{\mu \in U} \max_{\lambda} L(\lambda, \mu) \quad (3.28)
$$

**Slater’s Condition**

We can show strong duality between the primal, centralized problem and the Lagrangian dual using Slater’s condition. Slater’s condition states that for any convex optimization problem

$$
\begin{align*}
\min & \quad f_0(x) \\
\text{s.t.} & \quad f_i(x) \leq 0, \quad i = 1, \ldots, m \\
& \quad Ax = b
\end{align*}
$$

strong duality holds so long as $\exists x \in \text{relint}(D)$, where $D$ is the domain, such that

$$
\begin{align*}
& \quad f_i(x) < 0, \quad i = 1, \ldots, m \quad \text{and} \quad Ax = b
\end{align*}
$$

We can apply Slater’s condition to show strong duality. Because the cost function is convex, and all of our constraints are affine constraints, we can apply the weak form of Slater’s condition wherein affine constraints do not need to hold with strict inequality [36]. This means that strong duality holds so long as there exists $x$ such
that all constraints 3.5 through 3.11 are satisfied. The only possibility of infeasibility in the model is that $b_{\text{max}}$ is set to be a hard constraint, instead of allowing for energy spillage, as described in Section 2.3.1. This can be handled by adding an extra variable that absorbs any spilled energy and an associated cost of spillage. Thus, there is no duality gap between the primal and dual problems: $p^* = d^*$.

### 3.3 Subgradient Method

In order to update the price vector $\lambda$, the maximization over $\lambda$ can be solved using a subgradient method. A subgradient of a function $f$ at $x$ is any vector $s$ that satisfies

$$f(y) \geq f(x) + s^T(y - x)$$

The subgradient method is an iterative approach where $\lambda$ is updated every step by:

$$\lambda^{k+1} = \lambda^k - \alpha_k s^k$$

where $\alpha$ is an appropriate step size. There are many well known convergence results for the subgradient method with varying step sizes. [38] The general convergence proof is easy to show given two assumptions: that the norm of the subgradients are bounded, ie $||s^k||_2 \leq G$ and there is a bound on the distance of the initial point to the optimal set: $R \geq ||\lambda^1 - \lambda^*||_2$, and results in the inequality

$$f_{\text{best}} - f^* \leq \frac{R^2 + G^2 \sum_{i=1}^k \alpha_i^2}{2 \sum_{i=1}^k \alpha_i}$$

This result tells us multiple things. First, we can find a step size in which the suboptimality bound is reduced to zero (nonsummable diminishing step lengths) [38]. However, it also shows us that the subgradient method can be very slow.

To find the subgradient of $g(\lambda)$, we begin with our original definition 3.14 with
optimal values $v_i^*, k_i^*, f_{ij}^*$:

$$g(\lambda) = \min_{k_{ji}, f_{ij}, v_i} \sum_{i \in N} [x_i^T Q_i x_i + \sum_{j \in O(i)} f_{ij}^T \hat{R}_{ij} f_{ij} + \sum_{j \in I(i)} \{k_{ji}^T \hat{R}_{ji} k_{ji} + \lambda_{ji}^T (k_{ji} - f_{ji})\}]$$

The subdifferential of this in relation to $\lambda_{ji}$ is

$$s_{ji}(t) = k_{ji}^* - f_{ji}^* \quad (3.30)$$

Notice that each subgradient $s_{ji}(t)$ depends only on local variable $k_{ji}(t)$ and nearest neighbor decisions $f_{ji}(t)$. Therefore we can update each $\lambda_{ji}$ at every link using only information passed from nearest neighbors:

$$\lambda_{ji}^{k+1} = \lambda_{ji}^k - \alpha k s_{ji}^k$$

### 3.3.1 Sub-optimality bounds

Here we verify the assumption that function $g(\lambda)$ is Lipschitz continuous, or that the norm of our subdifferential $s$ is bounded. All flow variables are bounded by a maximum transmission flow value $c$, so both $k_{ji}$ and $f_{ji}$ are both bounded above and below by $c$. Therefore should all flows be maximal and with $k_{ji}$ and $f_{ji}$ opposite, each subdifferential $s_{ji}$ would be at most $2c$. The maximal bound of the 2-norm of $s$ would therefore be:

$$\|s^k\|_2 \leq \sqrt{\sum_{t=1}^T \sum_{i \in N} (2c)^2} \leq \sqrt{4c^2TN} \leq 2c\sqrt{TN}$$

Therefore we have

$$G = 2c\sqrt{TN} \quad (3.31)$$
We can use Equation 3.31 and Equation 3.29 to bound sub-optimality with varying step size choices. For example, using fixed step size $\alpha$, equation 3.29 becomes

$$f^k_{\text{best}} - f^* \leq \frac{R^2 + G^2 \alpha^2 k}{2\alpha k} \quad (3.32)$$

which converges to $G^2\alpha/2$ as $k$ goes to infinity [38].

### Algorithm 1 Distributed Optimization with Sub-Gradient Method

**Input:** $Q, R$

1: **Initialize** Set $z = \lambda$, $t_1 = 1$

2: **for** $k \geq 1$ **do**

3: \[
\hat{f}_{ij} \forall j \in O(i), \hat{k}_{j|i} \forall j \in I(i), \hat{v}_i \\
= \arg\min[x_i^T Q_i x_i + f_{ij}^T \bar{R}_{ij} f_{ij} + k_{ji}^T \bar{R}_{ji} k_{ji} + \lambda_{ji}^T k_{ji} - \lambda_{ij}^T f_{ij}]
\]

subject to:

$x_i = \bar{1} x_0 + L(v_i - d_i)$

$b_i = \bar{1} b_0 + L(v_i + s_i + \sum_{j \in I(i)} k_{ji} - \sum_{j \in O(i)} f_{ij})$

$v_i \geq 0$

$x_i \leq 0$

$b_i \geq 0$

$b_i \leq b_{\text{max}}$

$-c \leq f_{ij} \leq c$

4: $s_{ji}^k = \hat{k}_{ji} - \hat{f}_{ji} \forall j \in I(i)$

5: $\lambda_{ji}^{k+1} = \lambda_{ji}^k + \alpha_k s_{ji}^k$

Though the subgradient method is known to converge for appropriate step-size, it is impractically slow. Finding the optimal step size requires knowledge of the optimal solution ahead of time, which is unavailable in practice. And even in those situations, many examples still display slow convergence rates [38].

### 3.4 Fast Gradient Method

The subgradient method shows that the CLQR problem can be decomposed optimally. However, the subgradient method can converge very slowly. We can instead implement a fast gradient method for the price update. The first fast, or “accelerated” gradient method was presented by Nesterov in 1983, and has been extended and generalized [39–42]. Fast gradient methods achieve convergence rates of $O(1/k^2)$ ver-
sus $O(1/k)$ for the standard gradient method. Fast gradient methods can have been
applied to some dual decomposition problems; Beck et al develop a distributed fast
gradient method for solving the dual Network Utility Maximization (NUM) problem,
and Giselsson develops a fully distributed algorithm for distributed model predictive
control [43,44].

Some preliminary assumptions in all fast gradient methods is that the function
$f$ to be minimized is convex, differentiable, and has a Lipschitz continuous gradient
with constant $L$:
$$
\|\nabla f(x_1) - \nabla f(x_2)\|_2 \leq L \|x_1 - x_2\|_2
$$
Here we show that our primal function has a Lipschitz continuous gradient using the
strong convexity of the primal problem. We follow a similar methodology to Beck et
al. [43]. Recall from Section 3.2.1 that the centralized problem can be formulated as
a quadratic program
$$
f(u) = \min_u T x_0^T + u H u + 2 \vec{1} x_0 F u 
\tag{3.33}
$$
with
$$
u(t) = [\vec{v}_1(t) \ldots \vec{v}_N(t) \ f_{ij}(t)]^T
$$
$$
H = \begin{bmatrix}
L^T L \\
\vdots \\
L^T L \\
R
\end{bmatrix}
$$

This can be written in terms the additional decision variables $f_{ji} = k_{ji}$. Define
global decision vector
$$
u = [\vec{v}_1 \ldots \vec{v}_N, \ f_{ij} \ldots f_{Nj}, \ k_{j1} \ldots k_{jN}]^T
$$
The constraints that we are forming the Lagrangian with are \( k_{ji}(t) = f_{ji}(t) \) for all \( j \in I(i) \). To write this in compact form, define

\[
C = \begin{bmatrix}
0 & -I & I \\
-I & I
\end{bmatrix}
\]

where \( I \) are identity matrices with dimensions \((m \times T)\). Then the equality constraint \( k_{ji}(t) = f_{ji}(t) \) can be equivalently written as

\[
Cu = 0
\]

We can now write the corresponding Lagrangian dual function as

\[
g(\lambda) = \min_u f(u) + \lambda^\top (Cu)
\]

The conjugate \( f^* \) of a function \( f \) is given by:

\[
f^*(y) = \sup_{u \in \text{dom}(f)} (y^\top u - f(u))
\]

Notice that the Lagrangian dual \( g(\lambda) \) is equal to

\[
g(\lambda) = -\max_u (-C\lambda^\top u - f(u))
\]

\[
= -f^*(-C\lambda)
\]

Therefore, \( g(\lambda) \) is Lipschitz continuous by an equivalence between strong convexity in the primal function and properties of its conjugate, proposed in [45], which states
that for a proper, lower semi-continuous, convex function $f : \mathbb{R}^n \to \mathbb{R}$ and a value $\sigma$ the following two properties are equivalent:

(a) $f^*$ is strongly convex with constant $\sigma$
(b) $f$ is differentiable and $\nabla f$ is Lipschitz continuous with constant $1/\sigma$

We can use this to find a Lipschitz constant for $g(\lambda)$. Beck et al. shows that under these conditions the function $g$ has a Lipschitz continuous gradient with constant $\|C\|^2_2/\lambda_{\min}(H)$, and that the primal problem converges from $u \to u^*$ at a rate of $\mathcal{O}(1/k)$ [43]. Gisselson et. al shows that for a strongly convex cost function, the tightest lower bound on the Lipschitz constant is $L \succeq CH^{-1}C^\top$. Thus we can implement Nesterov’s optimal gradient method, shown in Algorithm 2. The tradeoff

Algorithm 2 Distributed Optimization with Fast Gradient Method

**Input:** $L \leftarrow$ Lipschitz constant of $\nabla g(\lambda)$

1. **Initialize** Set $z = \lambda$, $t_1 = 1$
2. **for** $k \geq 1$ **do**
   3. $\hat{f}_{ij} \forall j \in O(i), \hat{k}_{ji} \forall j \in I(i), \hat{v}_i$
      = arg min$[x_i^\top Q_i x_i + f_{ij}^\top \hat{R}_{ij} f_{ij} + k_{ji}^\top \hat{R}_{ji} k_{ji} + \lambda_{ji}^\top k_{ji} - \lambda_{ij}^\top f_{ij})]$ subject to:
         $x_i = \bar{x} x_0 + L(v_i - d_i)$
         $b_i = \bar{b} b_0 + L(v_i + s_i + \sum_{j \in I(i)} k_{ji} - \sum_{j \in O(i)} f_{ij})$
         $v_i \geq 0$
         $x_i \leq 0$
         $b_i \geq 0$
         $b_i \leq b_{\text{max}}$
         $\bar{v}_i \leq f_{ij} \leq \bar{c}$
   4. $u = [\bar{v}_1 \ldots \bar{v}_N, f_{1j} \ldots f_{Nj}, k_{j1} \ldots k_{jN}]^\top$
   5. $\lambda^k = z^k + L^{-1}(Cu)$
   6. $t^{k+1} = \frac{1}{1+\sqrt{1+4(t^k)^2}}$
   7. $z^{k+1} = \lambda^k + \frac{t^{k-1}}{t^{k+1}}(\lambda^k - \lambda^{k-1})$

with using the fast version of the subgradient method is that the Lipschitz gradient is calculated centrally, and $\lambda$ updates are done centrally. This allows for the maintenance of propriety information at each microgrid, but requires a centralized hub to perform the price calculations. There are techniques to do the dual update in a fully distributed method should the network conditions require it [43,44].
3.5 Dual Decomposition Example

Here an example is provided to show the output of the subgradient method. The method is applied to a five node graph that is connected by six links (see Figure 3-2). The optimization was done over twelve time steps. Supply and demand were generated randomly for each time point at each node from a uniform distribution between zero and 1. All initial values were set to zero. In this example we used constant step size $\alpha = .01$ and terminated the program when the subgradient $s_{ji} < .001$ for all links. Using a subgradient method, convergence took 19,250 iterations versus 431 using the fast gradient method.

Figure 3-2: Example Microgrid Network
Figure 3-3: Flow decisions for each node using centralized and dual decomposition methods
Chapter 4

Dual Decomposition of the LQR Problem Under Network Flow Constraints

As discussed in Chapter 2, the constraints on the model explored in Chapters 2 and 3 destroy the structure of the LQR problem. In this chapter we focus on a more general class of problems. We study a network comprised of local LQR problems whose decisions are subject to network flow constraints. We use dual decomposition of the network flow constraints to develop a fully distributed solution that relies on message passing between nearest neighbor nodes. We then develop a model predictive control algorithm to compute an approximate distributed policy using only local state information under uncertain disturbances.

Various techniques for decomposition and coordination for dynamic systems have been explored in the literature. Cohen develops a generalized iterative algorithm for the decentralization of dynamic optimal control problems, extending upon Takahara’s algorithm [46]. Martensson et al. provide a distributed solution for local linear feedback controller using an iterative gradient method [47]. The method of dual decomposition has also been used extensively to coordinate dynamic systems coupled by state and input constraints using model predictive control [44, 48, 49].

In this chapter, we apply a dual decomposition technique to a specific case of
the decentralized linear quadratic regulator. The work in this chapter extends two very well known theories. The first is the well known Linear Quadratic Regulator problem, as described in Chapter 2. Recall that for quadratic cost function $J(x_t, u_t)$

$$J(x_t, u_t) = \min_u E_{w_t} \{x(T)^T Q x(T) + \sum_{t=0}^{T-1} x(t)^T Q x(t) + u(t)^T R u(t)\}$$

subject to: $x(t+1) = A x(t) + B u(t) + w(t)$

the optimal solution is given by

$$\mu_t(x_t) = L_t x_t$$

where $L_t$ is determined by solving the discrete time Ricatti equation

$$L_t = -(B^T K_{t+1} B + R)^{-1} B^T K_{t+1} A$$

with $K$ defined as

$$K_t = Q + A^T K_{t+1} A - A^T K_{t+1} B (R + B^T K_{t+1} B)^{-1} B^T K_{t+1} A$$

Rantzer shows that dual decomposition can be applied to decompose dynamic feedback systems with separable quadratic costs whose states are coupled through matrix $A$. However, the optimal policy still depends on full information of all states $x$ [50].

The second theory we draw upon is the dual decomposition of the single network commodity flow problem. In the single commodity network flow problem, the goal is to choose the optimal flow across some network $G = (N, E)$ subject to flow conservation constraints. The single network commodity flow problem can be summarized as follows: Define $f_j$ as the flow on arc $j$, where $f_j > 0$ implies flow is in the direction of the link. Then the network flow conservation requirement can be written as $C f + d = 0$ where $C$ is the incidence matrix of graph $G$. Given some separable cost function $\phi(f) = \sum_{j=1}^n \phi_j(f_j)$, with $\phi_j: \mathbb{R} \to \mathbb{R}$ being the strictly convex cost function
for arc $j$, the Lagrangian dual function can be formulated with no duality gap. Solving the network flow optimization via the dual results in the optimal solution being given as a function of the optimal potential difference across each link $j$, where the potential refers to the difference between the dual variables of the end node and start node of link $j$ [51].

There are some nice features of both the unconstrained LQR problem and dual decomposition of the network flow problem. First, the optimal solution to the LQR problem can be calculated ahead of time as a function of current state $x_i$, making it a very useful solution for handling uncertainty online. The advantage of the network flow problem is that it elegantly decomposes to have strong duality while only relying on nearest neighbor communication. In this section, the theories underlying these two problems are leveraged to find a distributed solution to a set of networked LQR problems, where the local states at each node depend on decisions that are subject to network flow constraints.

### 4.1 The Model

We are given some network represented by a connected, directed graph $G = (N, E)$ with known topology of $n$ nodes and $m$ edges. Each node $n$ can have multiple local state variables $x_i \in \mathbb{R}^{z_i}$. At each time point, nodes must choose the amount of commodity $d_i$ to withdraw from or add to the network. State $x_i$ represents a buffer, allowing flexibility in the amount of resource $d_i$ to supply or demand from the network at each time. The cost function is quadratic in local states $x_i(t)$ as well as in flow decisions $u = [f_{ij}]^\top$ for all $(i, j) \in E$. Define $x = [x_i]^\top$ for all $i \in N$. Then the cost function is given by

$$
\min_u \mathbb{E}[x(T)^\top Q x(T) + \sum_{t=1}^{T-1} x(t)^\top Q x(t) + u(t)^\top R u(t)]
$$

(4.1)
subject to:

\[ x_i(t+1) = A_i x_i(t) + \vec{1} d_i + w_i(t) \quad \forall i \]  
\[ \sum_{j \in I(i)} f_{ji}(t) - \sum_{j \in O(i)} f_{ij}(t) = d_i \quad \forall i \]  

It is possible to combine the network flow constraint in Equation 4.3 into the dynamics Equation 4.2:

\[ x_i(t+1) = A_i x_i(t) + \vec{1} \sum_{j \in I(i)} f_{ji}(t) - \vec{1} \sum_{j \in O(i)} f_{ij}(t) + w_i(t) \quad \forall i \]

Then the system wide dynamics can be written in terms of \( x \) and \( u \) as

\[ x(t+1) = A x(t) + \tilde{B} u(t) + w(t) \]

where \( \tilde{B} \) is given by the incidence matrix of the network

\[ \tilde{B}_{ij} = \begin{cases} 
-1, & \text{if i is the start node of the jth arc} \\
1, & \text{if i is the end node of the jth arc} \\
0, & \text{otherwise}
\end{cases} \]

with row \( i \) repeated \( z_i \) times for each local state \( x_i \). Assume that matrices \( Q \) and \( A \) are block diagonal with \( n \) blocks \( Q_i \) and \( A_i \) of size \([z_i \times z_i]\), and that \( Q \) is positive semidefinite and each block \( A_i \) is stable. Also assume \( R \) is a positive definite diagonal matrix. Just as in Chapter 2, the flow decisions are coupled across nodes. In order to decouple the problem we apply dual decomposition. Once again, define variables
\( k_{ji} = f_{ji} \); then the problem can be formulated as

\[
\min_u \mathbb{E}_{w_t} \left[ \sum_{i \in N} x_i(T)^\top Q x_i(T) + \sum_{t=0}^{T-1} (x_i^\top(t)Qx_i(t) + u(t)^\top \hat{R}u(t)) \right]
\]

subject to:

\[
x_i(t+1) = A_i x_i(t) + \sum_{j \in I(i)} k_{ji}(t) - \sum_{j \in O(i)} f_{ij}(t) + w_i(t) \quad \forall i \in N
\]

\[
k_{ji}(t) = f_{ji}(t) \quad \forall j \in I(i)
\]

The coupling in our decisions is now confined to the last constraint \( k_{ji}(t) = f_{ji}(t) \).

To eliminate the coupling constraint we formulate the Lagrangian dual problem:

\[
L(\lambda, u_i, f_{ij}, k_{ij}) = \max_{\lambda} \min_{u_i, f_{ij}, k_{ij}} \mathbb{E}_{w_t} \left\{ \sum_{i \in N} x_i(T)^\top Q x_i(T) + \sum_{t=0}^{T-1} (x_i^\top(t)Qx_i(t) + u(t)^\top \hat{R}u(t)) + \lambda_{ji}(t)^\top (k_{ji}(t) - f_{ji}(t)) \right\}
\]

Using results from Propositions 1 and 2, and define \( u_i(t) = [f_{ij}(t) \quad k_{ij}]^\top \), then we know that we can write the Lagrangian as

\[
L(\cdot) = \max_{\lambda} \min_{u_i, f_{ij} \forall j \in O(i), k_{ji} \forall j \in I(i)} \mathbb{E}_{w_t} \left\{ \sum_{i \in N} x_i(T)^\top Q x_i(T) + \sum_{t=0}^{T-1} (x_i^\top(t)Qx_i(t) + u_i^\top(t) \hat{R}u(t)) + \lambda_{ji}(t)^\top k_{ji}(t) - \lambda_{ij}(t)f_{ij}(t) \right\}
\]

where \( \hat{R} \) is again defined as \( \frac{1}{2}R \). Define vector

\[
\tilde{\lambda}_i(t) = [\lambda_{ji}(t) - \lambda_{ij}(t)]^\top
\]

Now we can write the optimization at each node only in terms of \( u_i \). Stage cost at each node \( i \) can be written as

\[
l_i(\bar{x}_i(t)) = x_i(t)^\top Qx_i(t) + u_i(t)^\top R_i u_i(t) + \tilde{\lambda}_i(t)^\top u_i(t)
\]

(4.6)
The goal is to minimize total cost

$$
\mathbb{E}\{\sum_{i \in N} l_i(\tilde{x}_i(T), u_i(T)) + \sum_{t=1}^{T-1} l_i(\tilde{x}_i(t), u_i(t), \tilde{\lambda}_i(t))\}
$$

subject to:

$$
x_i(t + 1) = A_i x_i(t) + B_i u_i(t) + w_i(t)
$$

Local state dynamics $B_i$ is defined as a matrix of $z_i$ repeated rows each composed of the nonzero entries of $\tilde{B}_i$.

### 4.2 Policy Generation

For brevity during this section we drop the subscripts $i$ because each optimization is being done independently at node $i$. Instead we write time as a subscript.

**Proposition 4** The optimal distributed control policy $\mu_t(x_t, \tilde{\lambda}_t^f)$ for the distributed LQR is linear in local state $x_t$ and future dual variables $\tilde{\lambda}_t^f$, given by

$$
\tilde{\lambda}_t^f = [\tilde{\lambda}_t, \tilde{\lambda}_{t+1}, \ldots, \tilde{\lambda}_T]^\top
$$

We can show that in general, the cost function is of the form

$$
J_t(x_t) = x_t^\top K_t x_t + \tilde{\lambda}_t^f \beta_t x_t + \tilde{\lambda}_t^f P_t \tilde{\lambda}_t^f + \sum_{\tau=t}^{T-1} \mathbb{E}[w_\tau K_{\tau+1} w_\tau]
$$

and optimal control policy $\mu_t^*$ is

$$
\mu_t^* = -(R + B^\top K_{t+1} B)^{-1}(B^\top K_{t+1} A x_t + \frac{\tilde{\lambda}_t}{2} + \frac{B^\top \beta_{t+1}^\top \tilde{\lambda}_{t+1}}{2})
$$
Proof By backwards induction:

\[
J_{T-1}(x_{T-1}) = \min_{u_{T-1}} \mathbb{E}[\tilde{x}_{T-1}^TQ\tilde{x}_{T-1} + u_{T-1}^TRu_{T-1} + \tilde{\lambda}_{T-1}^Tu_{T-1} + \\
(A\tilde{x}_{T-1} + Bu_{T-1} + w_{T-1})^TQ(A\tilde{x}_{T-1} + Bu_{T-1} + w_{T-1})]
\] (4.7)

Expanding the above equation, and assuming that the mean of our disturbances are zero, eliminate cross terms with \(w\):

\[
J_{T-1}(x_{T-1}) = \min_{u_{T-1}} \left\{ \tilde{x}_{T-1}^TQ\tilde{x}_{T-1} + u_{T-1}^TRu_{T-1} + \tilde{\lambda}_{T-1}^Tu_{T-1} + \right.
\]
\[
+ u_{T-1}^TB^TQBu_{T-1} + 2x_{T-1}^TA^TQBu_{T-1} + x_{T-1}^TA^TQAx_{T-1} + \\
+ \mathbb{E}[w_{T-1}Qw_{T-1}] \right\}
\]

Differentiating with respect to \(u_{T-1}\) and setting the derivative equal to zero, we find optimal value of \(u_{T-1}\):

\[
u_{T-1}^* = -(R + B^TQB)^{-1}(B^TQAx_{T-1} + \frac{\tilde{\lambda}_{T-1}}{2})
\]

Plugging this back in for \(J_{T-1}\):

\[
J_{T-1}(x_{T-1}) = x_{T-1}^TK_{T-1}x_{T-1} + \tilde{\lambda}_{T-1}^T(R + B^TQB)^{-1}B^TQAx_{T-1} - \\
- \frac{\tilde{\lambda}_{T-1}^T}{2} (R + B^TQB)^{-1} \frac{\tilde{\lambda}_{T-1}}{2} + \mathbb{E}[w_{T-1}Qw_{T-1}]
\]

and define vector \(\beta_{T-1} = -(R + B^TQB)^{-1}B^TQA\). Then

\[
J_{T-1}(x_{T-1}) = x_{T-1}^TK_{T-1}x_{T-1} + \tilde{\lambda}_{T-1}^T\beta_{T-1}x_{T-1} - \\
- \frac{\tilde{\lambda}_{T-1}^T}{2} (R + B^TQB)^{-1} \frac{\tilde{\lambda}_{T-1}}{2} + \mathbb{E}[w_{T-1}Qw_{T-1}]
\]
Continuing induction, we have

\[
J_{T-2}(x_{T-2}) = \min_{u_{T-1}} \mathbb{E}[x_{T-2}^T Q x_{T-2} + u_{T-2}^T R u_{T-2} + \tilde{\lambda}_{T-2} \cdot u_{T-2}
+ (A \tilde{x}_{T-2} + B_{T-2} u_{T-2} + w_{T-2})^T K_{T-1} (A x_{T-2} + B_{T-2} u_{T-2} + w_{T-2})^T
+ \tilde{\lambda}_{T-1} \beta_{T-1} (A x_{T-2} + B_{T-2} u_{T-2} + w_{T-2})^T
- \frac{\tilde{\lambda}_{T-1}}{2} (R + B^T Q B)^{-1} \frac{\tilde{\lambda}_{T-1}}{2} + w_{T-1} Q w_{T-1}]
\] (4.8)

Taking the derivative over \(u_{T-2}\) we find

\[
u_{T-2}^* = -(R + B^T K_{T-1} B)^{-1} (B^T K_{T-1} A \tilde{x}_{T-2} + \frac{\tilde{\lambda}_{T-2}}{2} + \frac{B^T \beta_{T-1} \tilde{\lambda}_{T-1}}{2})
\]

Define

\[
\beta_{T-2} = \begin{bmatrix}
-(R + B^T K_{T-1} B)^{-1} B^T K_{T-1} A \\
- \beta_{T-1} (B (R + B^T K_{T-1} B)^{-1} B^T K_{T-1} - I) A
\end{bmatrix}
\]

and \([(T - t) \times 1]\) vector of future prices \(\tilde{\lambda}'_t \subset \tilde{\lambda}\)

\[
\tilde{\lambda}'_t = [\tilde{\lambda}_t, \tilde{\lambda}_{t+1}, \ldots, \tilde{\lambda}_T]^T
\]

Then the value function at \(T - 2\) can be written

\[
J_{T-2}(x_{T-2}) = x_{T-2}^T K_{T-2} x_{T-2} + \tilde{\lambda}'_{T-2} \beta_{T-2} x_{T-2}
+ \tilde{\lambda}'_{T-2} P_{T-2} \lambda_{T-2} + \sum_{\tau=T-2}^{T-1} \mathbb{E}[w_{\tau} K_{\tau+1} w_{\tau}]
\]

For arbitrary time \(t\):

\[
J_t(x_t) = x_t^T K_t x_t + \tilde{\lambda}'_t \beta_t x_t + \tilde{\lambda}'_t P_t \tilde{\lambda}'_t + \sum_{\tau=t}^{T-1} \mathbb{E}[w_{\tau} K_{\tau+1} w_{\tau}]
\]
Then for \( J_{t-1} \):

\[
J_{t-1}(x_{t-1}) = \min_{u_{t-1}} \mathbb{E}[x_{t-1}^T Q x_{t-1} + u_{t-1}^T R u_{t-1} + \tilde{\lambda}_{t-1}^T u_{t-1} + \begin{pmatrix} (A x_{t-1} + B u_{t-1} + w_{t-1})^T K_t (A x_{t-1} + B u_{t-1} + w_{t-1}) \\ + \tilde{\lambda}_{t}^T \beta_t (A x_{t-1} + B u_{t-1} + w_{t-1}) \\ + \tilde{\lambda}_{t}^T P_t \tilde{\lambda}_{t} + \sum_{\tau=t}^{T-1} \mathbb{E}[w_{\tau} K_{\tau+1} w_{\tau}] \end{pmatrix}]
\]

Taking the derivative with respect to \( u_{t-1} \) and setting equal to zero, we find

\[
u_{t-1}^* = -(R + B^T K_t B)^{-1}(B^T K_t A x_{t-1} + 1/2(\tilde{\lambda}_{t-1} + B^T \beta_t^T \tilde{\lambda}_{t}))
\]

Plugging back into the value function with

\[
\beta_t = \begin{bmatrix} -(R + B^T K_t B)^{-1} B^T K_t A \\ -\beta_t (B(R + B^T K_t B)^{-1} B^T K_t - I) A \end{bmatrix}
\]

and

\[
\tilde{\lambda}_{t-1}^f = [\tilde{\lambda}_{t-1}, \tilde{\lambda}_t, \ldots, \tilde{\lambda}_T]^T
\]

we can re-write the cost to go as

\[
J_{t-1}(x_{t-1}) = x_{t-1}^T K_{t-1} x_{t-1} + \tilde{\lambda}_{t-1}^T \beta_{t-1} x_{t-1} + \tilde{\lambda}_{t-1}^T P_{t-1} \tilde{\lambda}_{t-1} + \sum_{\tau=t}^{T-1} \mathbb{E}[w_{\tau}^T K_{\tau+1} w_{\tau}]
\]

which is the same form as we started with.

### 4.3 Price update and Receding Horizon Control

The optimal distributed decision depends on the values of current and future dual variables. For every iteration \( k \), we approximate the values of dual variables \( \tilde{\lambda} \) by performing a forward propagation of our system under policy \( \mu^k(\tilde{\lambda}) \) and the expected
value of disturbances $w$. This is shown in Algorithm 3. To update the dual vector $\tilde{\lambda}_i(t)$, a gradient method can be implemented. Recall from equation (4.6) that the gradient over the nodal dual variable $\lambda_{ji}$ is given by $s_{ji} = k_{ji} - f_{ji}$. This update must happen at every node, relying on the communication from neighbors of their flow decisions $f_{ji}$:

$$\lambda_{ji}^{k+1} = \lambda_{ji}^k - \alpha_ks_{ji}^k$$

After calculating $\lambda_{ji}(t)$ at each node for all $t$, $\tilde{\lambda}_i(t)$ can be formed by communicating values of $\lambda$'s between the nodes' neighbors. Convergence of the algorithm follows from the same subgradient results shown in Section 3.3.

For deterministic problems, the forward propagation of the system is exact, and the policy provides the exact optimal solution. Under uncertainty, the estimation of the future price vector is an approximation given predicted disturbance $E[w(t)]$. Because the future prices are estimated using current local state $x_i$, the algorithm is implemented as a receding horizon control algorithm. This means that the policy is applied to the first decision, then the calculation is done again once the disturbance at that time point has been realized. The advantage of the algorithm over conventional model predictive control solutions using dual decomposition is that the minimization problem has been calculated offline, so no minimization has to take place at every iteration.

At every stage, there is error accruing for two reasons. The first is the error in $\lambda_i$ given our decision to terminate the algorithm when $\max(s_{ji}) < \epsilon$. The second is during the forward propagation of the system dynamics. In order to generate the future price estimates, the future states $x_i$ must also be estimated by propagating the system forward $x_i(t+1) = Ax_i(t) + Bu_i(t)$ (because $E[w_i(t)] = 0$). For each time point $t$ any variance in the disturbance $w_i$ will be multiplied by matrix $A$. Thus $A$ is constrained to be a stable matrix.
Algorithm 3 Distributed LQR

Input: $\tilde{\lambda}_i, \tilde{x}_i(1)$

1. **Initialize** Set $\tilde{\lambda}_i = \tilde{0}$
   
   $K_i(T) = Q,$
   $\beta_i(T - 1) = -(R_i + B_i^T Q B_i)^{-1} B_i^T Q A$

2. **for** $t = T : -1 : \tau$ **do**

3. $K_i(t - 1) = A_i^T (K_i(t) - K_i(t)B_i(B_i^T K_i(t)B_i + R_i)^{-1} B_i^T K_i(t))A_i + Q_i$

4. **for** $t = T - 1 : -1 : \tau$ **do**

5. $\beta_i(t - 1) = \begin{bmatrix} -(R_i + B_i^T K_i(t)B_i)^{-1} B_i^T K_i(t)A_i \\ -\beta_i(t)(B_i(R_i + B_i^T K_i(t)B_i)^{-1} B_i^T K_i(t) - I)A_i \end{bmatrix}$

6. **for** $k \geq 1$ **do**

7. **for** $t = \tau : T - 1$ **do**

8. $u_i(t) = -(R_i + B_i^T K_i(t + 1)B_i)^{-1}(B_i^T K_i(t + 1)Ax_i(t) + 1/2(\tilde{\lambda}_i(t) + B_i^T \beta_i(t + 1)^T \tilde{\lambda}_i^T(t)))$

9. $x_i(t + 1) = A_i x_i(t) + B_i u_i(t)$

10. $s_k^i = k_{ji} - f_{ji}$ where $k_{ji} \in u_i$ and $f_{ji} \in u_j \forall j \in I(i)$

11. $\lambda_{ji}^{k+1} = \lambda_{ji}^k + \alpha_k s_{ji}$

12. $\tilde{\lambda}_i = [\lambda_{ij} - \lambda_{ji}]^T$

4.4 Performance

At each stage the control policy attempts to minimize cost in spite of disturbances. To be optimal, the decision of how much of the resource to add or remove from each buffer requires information about all buffer state $x$. In order to assess the performance of the algorithm, we can compare the decisions generated from Algorithm 3 to the optimal decisions made using full information for all states $x$, given by the LQR solution: $\mu_t(x_t) = L_t x_t$. Here we present an example of the distributed algorithm in comparison to the centralized optimal solution.

The network that is optimized over is shown in Figure 4-1. In the example, there are four nodes and four links. $Q_i$ and $R_i$ are the identity matrices. Incidence matrix for the graph is given by

$$\tilde{B} = \begin{bmatrix} -1 & 0 & 0 & 1 \\ 1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}$$
Each node has two states, and

\[ A_i = \begin{bmatrix} .25 & .30 \\ .40 & .32 \end{bmatrix} \]

Therefore each matrix \( B_i \) is as follows:

\[ B_1 = \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \end{bmatrix}, \quad B_3 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \quad B_4 = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \]

Initial states \( x_o \) and offsets \( w_i(t) \) at each node are generated from a uniform distribution between \([-0.5, 0.5]\). Results comparing the distributed and centralized solutions are compared in Figure 4-2.
Figure 4-2: State evolution under optimal (centralized) and distributed policies.
Chapter 5

Conclusions and Future Work

5.1 Overview

As discussed in Chapter 1, electricity access is a global issue that India is particularly affected by, with an estimated 240 million people without access. Electrification through conventional generation and grid extension suffers from many drawbacks. Extending the central grid to villages that are far from existing power plants is very costly; extending the grid just 5 to 25 km from the nearest existing lines is not cost competitive. Furthermore, the central government has made electrification through renewable sources a priority. This has brought down costs of centralized solar projects significantly. Unfortunately, on-site generation continues to suffer from a lack of financing, leading to sporadic implementation at small scales. Any technological advantage that would provide for lower upfront capital costs would allow for a more competitive energy landscape in which small microgrid operators may be able to compete more effectively with subsidized energy prices.

5.2 Value of Interconnection

In Chapter 2, a model was developed that would allow for power to be shared across distinct, individual microgrids. Costs were associated with unmet, backlogged demand, and the cost of transmission between microgrids. The decisions involve opti-
mizing both the local demand schedule as well as the amount of energy to transmit between nodes over a finite team period. The model is deterministic with some supply and demand forecast, and can be solved centrally as a quadratic program.

Using this model and supply and demand data gathered in India, we showed that storage capacity could be removed from some microgrids without significantly altering the overall reliability of the microgrids. Simulations were done to emulate a 25 microgrid network being powered by a 250W solar panel. During months when there was only critical demand on the system, up to %60 of the storage could be removed from the network with only a 2% decrease in overall reliability. Over the whole year, %40 percent of storage could be removed with just a 3% decrease in reliability. However, there are tradeoffs to removing local storage. These reliability estimates depend on the cost of transmission relative to the cost of serving unmet demand, which is directly related to the distance between the microgrids. If storage is too far from the microgrid it needs to serve, costs of transmission could outweigh cost of serving demand. In Section 2.3, supply side uncertainty was introduced. We found that the deterministic model outperforms a system without interconnection even with supply side variability.

5.3 Decomposability of the Model

Chapter 2 proposed a centralized control system for microgrid interconnection. Often this may be infeasible in practice, because of limitations in centralized resources and communication structures, and prioritization in retaining proprietary data. In Chapter 3, the model is decomposed to allow for local autonomy and decision making of microgrids. We showed that using a dual decomposition technique, we can separate the optimization so that the microgrids can maintain proprietary supply and demand forecasts and compute scheduling optimizations locally. The outputs of these local optimizations are used to update price variables throughout the network, until an optimal global solution is reached. We show that there is no duality gap between the distributed solution and the centralized solution, and that only local communi-
cation between nearest neighbors is required to facilitate the price variable updates. This means that interconnection could be implemented in a completely decentralized fashion.

5.4 Extension to the Distributed LQR

In Chapter 4, the ideas that were generated in formulating the interconnection model are extended to a more general setting. We focused on a specific class of problems in which a quadratic cost function is subject to linear dynamics whose decisions depend on network flow constraints. There are many general methods for decomposing dynamic systems in order to apply local optimization. For this problem, we applied dual decomposition to the network flow conservation constraints that coupled the systems. Both the LQR problem and the dual decomposition of network flow constraints have elegant solutions. By combining these two bodies of work, we generated an approximately optimal, fully distributed policy for each state only in terms of local state variables \( x_i(t) \) and price variables \( \tilde{\lambda} \), which only involves message passing between nearest neighbors. The algorithm presented in Algorithm 3 updates the policy \( \mu_i(x_i(t)) \) given some estimates of future price vectors \( \tilde{\lambda}_f \). The algorithm relies on approximations depending on current states \( x_i(t) \) and is run as a receding horizon control problem.

5.5 Future Work

Going forward, it would be useful to simulate the model under a larger variety of simulation parameters. It would be interesting to assess sensitivity and performance under a variety of conditions. For example, simulating under homogeneous supply and demand gives a conservative estimate of the value of interconnection; in circumstances with more variable supply and demand between nodes interconnection could provide additional value. A full Monte-Carlo simulation could be generated in order to find the affects under a variety of different active constraint circumstances. For
implementation, parameter choices would have to be made based on the specific net-
work and context, but having a knowledge of sensitivity to different parameters would
assist in the decision making process. Furthermore, if implementation ever happened,
it would be interesting to see feedback on the model itself. There may be practical
conditions that would change the way that the model should be designed for use in
the field.

In general, there are interesting questions that could tackled with the knowledge
that microgrid interconnection potentially has value to offset storage capacity needs.
In Section 2.3, we discussed the impact that storage siting has on microgrids with
similar supply and demand profiles. A potential research question would be to thor-
oughly explore the storage placement question in a more general sense. It would also
be interesting to explore network effects in greater detail. For example, although the
model incorporates congestion constraints, they were often not active at the supply
and demand scale that we are optimizing for. However, it would be important to
assess the effects of congestion on more general microgrid networks, and how they
might impact the fragility of the system.

In the future, uncertainty should be addressed directly in the model. We showed
the the deterministic model is still an improvement over no interconnection, even with
supply side variability. However, this could be made optimal by including variability
in the policy making process. It is possible that the algorithm developed in Chapter
4 could be modified and extended to include constraints. This could possibly be
done by another level of dual decomposition. Finally, Algorithm 3 could be modified
to include decision variables that are not only flow decisions. This would involve
including all decisions in variable $u_i$, corresponding to adding zero terms into price
vector $\tilde{\lambda}_i$. This would generalize the types of optimizations covered by this Algorithm
beyond problems that only depend on total inflow at node $i$. 

76
5.6 Final Remarks

This work assessed the value of connecting independent, solar powered microgrids in India into a larger network. The results show that there is a potential to gain value by interconnecting microgrids in India, and that having storage at every microgrid may not be the best use of expensive resources. Instead, connections between microgrids allow the storage capacity to be shared while capitalizing on slight variations between the supply and demand at local microgrids. Given the relatively low cost of implementing additional short distance transmission versus providing storage capacity, microgrid interconnection has the potential to lower costs of providing energy to those who reside in off-grid areas in India.
Bibliography


proach,” *IEEE Transactions on Automatic Control*, vol. 52, no. 10, pp. 1826–


