

California Solar Initiative

**RD&D** ■ Research, Development, Demonstration  
■ and Deployment Program



Final Project Report:  
**Advanced Grid-Interactive  
Storage**

Grantee:  
**Solar City**

April 2014



***[www.CalSolarResearch.ca.gov](http://www.CalSolarResearch.ca.gov)***

## PREPARED BY



3055 Clearview Way  
San Mateo, CA 94402

### Principal Investigator:

Eric Carlson  
ecarlson@solarcity.com  
650-963-5105

### Project Partners:

Tesla Motors  
UC Berkeley Renewable and Appropriate  
Energy Laboratory  
UC Berkeley Energy and Resources Group and  
Mechanical Engineering  
Pacific Gas & Electric

## PREPARED FOR

### California Public Utilities Commission

California Solar Initiative: Research, Development, Demonstration, and Deployment Program

## CSI RD&D PROGRAM MANAGER



### Program Manager:

Ann Peterson  
[Ann.Peterson@itron.com](mailto:Ann.Peterson@itron.com)

### Project Manager:

Stephan Barsun  
[Stephan.Barsun@itron.com](mailto:Stephan.Barsun@itron.com)

Additional information and links to project related documents can be found at

### DISCLAIMER

*"Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the CPUC, Itron, Inc. or the CSI RD&D Program."*

<http://www.calsolarresearch.ca.gov/Funded-Projects/>

# Preface

The goal of the California Solar Initiative (CSI) Research, Development, Demonstration, and Deployment (RD&D) Program is to foster a sustainable and self-supporting customer-sited solar market. To achieve this, the California Legislature authorized the California Public Utilities Commission (CPUC) to allocate **\$50 million** of the CSI budget to an RD&D program. Strategically, the RD&D program seeks to leverage cost-sharing funds from other state, federal and private research entities, and targets activities across these four stages:

- Grid integration, storage, and metering: 50-65%
- Production technologies: 10-25%
- Business development and deployment: 10-20%
- Integration of energy efficiency, demand response, and storage with photovoltaics (PV)

There are seven key principles that guide the CSI RD&D Program:

1. **Improve the economics of solar technologies** by reducing technology costs and increasing system performance;
2. **Focus on issues that directly benefit California**, and that may not be funded by others;
3. **Fill knowledge gaps** to enable successful, wide-scale deployment of solar distributed generation technologies;
4. **Overcome significant barriers** to technology adoption;
5. **Take advantage of California's wealth of data** from past, current, and future installations to fulfill the above;
6. **Provide bridge funding** to help promising solar technologies transition from a pre-commercial state to full commercial viability; and
7. **Support efforts to address the integration of distributed solar power into the grid** in order to maximize its value to California ratepayers.

For more information about the CSI RD&D Program, please visit the program web site at [www.calsolarresearch.ca.gov](http://www.calsolarresearch.ca.gov).

## Contents

Advanced Grid-Interactive Distributed PV and Storage .....	4
1 Executive Summary: The Benefits of Integrating Solar PV and Energy Storage .....	4
Introduction .....	5
Profiles .....	6
Summary - Objectives .....	7
2 Refine communication hardware and deploy prototype low-voltage grid interactive battery systems ..	9
2.1 Install a portfolio of Firm PV systems based on off-the-shelf, 48VDC hardware to begin immediate data collection to confirm product requirements .....	9
Introduction .....	9
Background .....	9
Subtask 1 – Select Candidate Sites .....	10
Subtask 2 – Selection of Power Electronics .....	15
Subtask 3 – Battery Cells and Balance of System .....	16
Subtask 4 – Audit, Design, and Permitting.....	16
Subtask 5 – System Installation and Inspection.....	17
Lessons Learned .....	17
Lead-Acid Battery Technology for Commercial and Industrial Applications .....	18
Conclusion.....	19
2.2 Internet-Based Storage Control Platform .....	20
Introduction .....	20
Communications System Architecture .....	20
Battery Module and BMS.....	21
Battery Inverter/Charger .....	21
PV Inverter .....	21
Site Gateway .....	21
SolarGuard Server .....	22
Communication Protocols.....	22
Gateway to PV Inverter/Charge Controller.....	22
Inverter/Charger to Battery Management System.....	23
Gateway to Inverter/Charger.....	23
Gateway to Server / Server-to-Server.....	23



Conclusion .....	23
3 Develop and Deploy a Grid-Interactive Li-ion Battery System .....	24
3.1 Integrate the Tesla Motors lithium-ion, high-voltage battery pack with a grid Interactive inverter, charger, and PV System .....	25
Project Goals .....	25
Summary of Developments .....	25
Technical Progress since CSI RD&D .....	30
Appendix – Product Specifications .....	31
3.2 Deploy integrated FirmPV / high-voltage storage systems .....	33
Introduction .....	33
Subtask 1 – Select Candidate Sites .....	33
Subtask 2 – Select Power Electronics.....	38
Subtask 3 – Audit, Design, and Permit Systems.....	38
Subtask 4 – System Installation and Inspection.....	40
Subtask 5 – Test Completed Systems .....	42
Conclusion.....	42
4 Monetization of the Value of PV-Paired Energy Storage.....	43
4.1 Advanced Energy Storage Market Survey.....	44
Introduction .....	44
Methodology.....	45
Survey Results .....	46
Conclusion.....	58
4.2 PV Variability Analysis .....	59
4.3 Optimal Rate Designs and ISO Services for Maximizing the Value of Combined PV and Storage....	87
4.4 Models and Control Strategies for Optimal Control of FirmPV Systems .....	143
4.6 Advanced Energy Storage – Financing Mechanisms.....	182
Introduction .....	182
Background .....	183
Discussion.....	183
Conclusion.....	187
Conclusions .....	189
Next Steps .....	190

# **Advanced Grid-Interactive Distributed PV and Storage**

## **1 Executive Summary: The Benefits of Integrating Solar PV and Energy Storage**

In recognition of the need to support the growth of battery storage technology and explore the economic benefits that storage and solar photovoltaic (PV) systems can deliver to the grid, the California Public Utilities Commission provided funding for this project through the California Solar Initiative (CSI) Research, Development, Demonstration and Deployment Program. This support provided a critical head start to the development of the technology and market for advanced energy storage.

The work under this grant was carried out by SolarCity, in partnership with Tesla Motors and the University of California (UC), Berkeley. Through SolarCity's experience deploying both residential and commercial PV systems, Tesla's expertise in battery chemistry and management, and UC Berkeley's expertise in energy markets and grid management, the project team was able to complete the development, deployment, and analysis of integrated PV and advanced battery systems.

Going in, the key objectives were to analyze the variability of PV systems and the corresponding needs for energy storage systems, develop a suitable advanced lithium ion technology, demonstrate the cost effective deployment of these systems, and identify the required market and regulatory changes needed to unlock the benefits to utilities, grids, and customers that these paired systems provide.

The initial phases of research conducted by UC Berkeley suggest that distributed PV has a very minimal impact on the grid at moderate penetrations and these impacts are decreased when PV assets are decentralized. In these moderate penetrations, PV actually improved the conditions on the grid, especially with respect to peak demand. In high penetrations the addition of an advanced energy storage system can provide more than enough complementary peak load reduction and voltage benefits. Paired storage and PV can provide other benefits to the grid and to utilities as well (including avoided cost of upgrades or added capacity) and to the host customer (backup power).

The development portion of the grant was focused on creating energy storage technology that could be easily integrated with PV installations. Tesla was able to implement the design feedback from SolarCity in the initial pilot research and to develop both commercial and residential pilot products. Key lessons were learned in regard to equipment design, system communications and control, ease of installation, compliance with NEC code, and regulatory barriers.

This battery technology was deployed successfully in a series of pilot installations on both commercial and residential sites. Despite some initial barriers to the interconnection process, all sites are fully operational. SolarCity gained valuable insights into installation constraints, the importance of education for permitting authorities, and customer feedback.

The insight gained during the course of this project, along with the inclusion of storage projects in the CPUC's Self Generation Incentive Program (SGIP) have enabled SolarCity and Tesla to bring both commercial and residential energy storage products to market. Over a hundred of these systems have already been installed across California and both companies are continuing to explore the value that energy storage can provide.

## Introduction

As electricity use forecasts continue to show increases in the near future<sup>1</sup> and distributed energy systems become a more common way to meet the elevated demand for power, the balance of the grid as a whole is gaining greater importance. A number of state policies are transforming the energy system in California to rely increasingly on renewable resources, including the state's Renewable Portfolio Standard, the California Solar Initiative, net-energy metering, as well as the state's cap and trade program pursuant to Assembly Bill 32. Collectively, these policies have put pressure on reduced reliance on the dirty generators that often supply energy demands during peak periods. Further, the increased cost of upgrading electrical infrastructure has brought the sustainability of the existing grid into question.

As the state reduces its reliance on conventional fossil-fueled generating resources and expands its use of renewable resources, there is growing interest in and a substantial need to identify ways to address the operational realities this changing mix of energy resources engenders, recognizing that many renewable resources are "as-available" in nature and subject to intermittency. This has been expressly recognized through the adoption of policies and programs to support broad development of storage resources,

---

<sup>1</sup> California Energy Demand 2010-2020 Adopted Forecast, <http://www.energy.ca.gov/2009publications/CEC-200-2009-012/CEC-200-2009-012-CMF.PDF>

including storage procurement mandate pursuant to AB 2514 and the inclusion of storage in the state's Self Generation Incentive Program (SGIP). Distributed energy storage has an enormous potential to provide some of the stability needed for grid infrastructure to continue supporting the energy needs of California, to reduce peak demand, and to provide valuable backup power to consumers as we shift to a resource portfolio comprised of ever increasing amounts of renewable capacity.

In recognition of the need to advance the technology and explore the applications of battery storage when paired with solar photovoltaic (PV) systems, the California Public Utilities Commission through the California Solar Initiative (CSI) Research, Development, Demonstration and Deployment Program agreed to fund a grant entitled "Advanced Grid-Interactive Distributed PV and Storage". The CSI incentive program enabled rapid growth of the solar market in California and saw the deployment of 1.34 Gigawatts of solar PV added to the California grid to date.<sup>3</sup> To bolster the future of the market that this program created, this grant supported research, development, and deployment of energy storage products paired with PV systems with the goal of exploring the benefits that these hybrid systems create for customers, utilities, and the grid. SolarCity signed a contract for the administration and management of this grant with Itron in January, 2011.

In order to advance this technology and the regulatory framework that supports it, SolarCity brought together Tesla Motors and The University of California, Berkeley to pilot several technologies, use cases, and controls systems, specifically looking at the role and operational impacts/opportunities of distributed storage deployed with customer-side of the meter photovoltaic system.

## Profiles

SolarCity is California's leading full service solar power provider for homeowners and businesses - a single source for engineering, design, financing, installation, monitoring, and support. Our company provides cost-effective financing that enables customers to eliminate the high upfront costs of deploying solar. SolarCity has more than 2,000 California employees based at 24 facilities around the state and has provided clean energy services to more than 26,000 California customers.

Tesla Motors is the global leader in performance electric vehicles and electric powertrain development and production. Headquartered in Palo Alto, CA and with its primary manufacturing facility in Fremont, CA, (the former NUMMI plant owned jointly by Toyota and GM) Tesla has produced over 30,000 vehicles to date, which have traveled cumulatively over 175MM miles. Additionally, the company has produced more than 1.5GWh of battery storage for the vehicle market. In 2009, under the grant funding provided for this project, Tesla's electric powertrain division began developing modular, scalable stationary storage battery systems based on the vehicle battery technology. Since that project completed, Tesla has brought multiple products to production and is the emerging leader in the stationary storage market.

The Energy and Resources Group (ERG) was established as an academic degree-granting program at UC Berkeley in 1973. ERG has become a unique interdisciplinary community of graduate students, core faculty, and over 100 affiliates and researchers from across the campus devoted to elucidating,

---

<sup>3</sup> Go Solar CA - Program Totals By Administrator, [http://www.californiasolarstatistics.ca.gov/reports/agency\\_stats/](http://www.californiasolarstatistics.ca.gov/reports/agency_stats/)

analyzing, and helping to solve some of the most pressing resource concerns of our age. A team of students and faculty from the ERG team brought expertise in power systems and integration and control of storage devices to this project. The overall objective of this team's role in the project was to determine the impact and value of coupling distributed storage with photovoltaic systems.

### Summary - Objectives

The focus of this project was initially to explore the ability for energy storage to “firm” intermittent generation from PV systems, reduce peak demand, and improve grid reliability. However, in addition to significant progress on developing those benefits, valuable progress was also made on the design, permitting, and installation of PV paired with storage as well. The original four goals that we started with were:

1. (Task 4) Establish extent of PV variability to determine energy storage system needs
2. (Task 2) Demonstrate a cost effective pairing of PV and energy storage for load shifting, demand reduction, and conventional ancillary services
3. (Task 3) Show that advanced, distributed, PV-coupled, grid-interactive storage solutions will reduce cost and carbon emissions and improve grid reliability and security
4. (Task 4) Identify utility retail and ISO wholesale rate structures, tariffs, and market mechanisms that will be necessary to bring combined PV and storage to new markets

In addition, the following goals were added:

1. (Task 3) Reduce the cost of installation (balance of system and labor)
2. (Task 3) Streamline the interconnection process for “firm” PV systems

Throughout the various tasks outlined in this report, the team sought to investigate the limitations of technology, policy, and market needs of energy storage and its benefits when paired with renewable facilities. The initial tasks focused on improving controls and implementation of existing technology. The next phases involved development and deployment of a new lithium ion energy storage product using Tesla's advanced battery technology. The next several phases involved exploring value and use cases for combined PV and energy storage systems. The full scope of this project provided valuable insight into what is needed to advance energy storage technology to a point where it can be effectively paired with PV generation to provide increased value and greater stability to the grid.

SolarCity applauds the CPUC for funding this project and many like it to not only increase our knowledge of how existing technologies work but enable new technologies to gain a foothold in the marketplace. The funding for basic research and product development was ultimately a significant contributor to reaching the tipping point for widespread commercialization of an energy storage and PV product. As a result of this research, SolarCity has installed over 100 systems and has more than 500 under contract through the SGIP program. Additionally, Tesla was able to deploy a 1MW system at their site in Fremont.

In the medium and longer term, the key issue will shift from assessing and vetting the technical potential of customer sited energy storage technologies to addressing the regulatory regime within which those systems are deployed. Currently, the number of use case which customer-sited storage can actively

address is somewhat limited not by the technical capabilities of these systems but rather by a regulatory framework that was not developed with distributed storage technologies on the customer side of the meter in mind. Interconnection requirements and rules around wholesale market access will need to be reviewed and substantially reformed to ensure that the significant additional services that customer sited energy storage solutions can provide is not stranded behind the meter.

The material discussed in this report applies only to work completed to meet the tasks and goals funded by this grant. This report does not discuss the more recent commercialization efforts by the research participants and the technology described in this report does not represent the current product and service offerings.

## **2 Refine communication hardware and deploy prototype low-voltage grid interactive battery systems**

The overall goal of Task 2 was to explore existing technology, both in terms of energy storage systems, and communications hardware, to develop the requirements for the following tasks. This initial research and prototyping phase was especially crucial to the development of the Tesla lithium-ion battery product and to the selection of accompanying communications, controls, and equipment.

### **2.1 Install a portfolio of Firm PV systems based on off-the-shelf, 48VDC hardware to begin immediate data collection to confirm product requirements**

#### **Introduction**

Task 2.1 was designed to provide refined engineering feedback for the Tesla battery development in Task 3 and operational data for the refinement of business practices in Task 4. The first phase of the installation pilot, SolarCity needed to find pilot sites to install the initial lead acid systems for testing. Initial selection of pilot sites was delayed due to difficulty of defining the value proposition to the end customer. Although there is a considerable amount of value that energy storage can provide to both end customers and to the grid, the tariffs and markets needed to capture this value do not currently exist. For residential customers, the defining value is the ability to power their home in the event of a grid outage.

In order to make it possible to both have the grid interactivity and the backup capabilities, an AC-coupled system configuration was chosen using the SMA Sunny Boy/Sunny Island platform. The next step was putting together preliminary designs based on this configuration that could be vetted for installation requirements, cost, permitting, and interconnection requirements.

Three pilot sites were selected from existing solar PV installations under SolarCity management. They were selected because the customers had a strong interest in the development of battery technology and showed flexibility to pilot new technologies. They were all selected from areas near SolarCity's headquarters in San Mateo for ease of installation and data gathering. The value of the backup energy that these systems provide was the main feature that was presented to these residential customers.

#### **Background**

The goal of this phase was primarily to set a baseline for the performance with existing lead acid technology. This baseline included ease and cost of installing energy storage equipment, power electronics, and balance of system, the performance of equipment, and limitations of current technology. In order to adequately assess the viability of li-ion grid interactive storage, some assessment of commonly used energy storage technology (primarily for battery backup) was required. The assessment in this phase of the project helped shape the li-ion product specifications for both residential and commercial applications.

Initially, six lead-acid pilot installations were planned comprising three residential sites and three commercial sites. The scope was narrowed to encompass residential sites only at this stage. Due to the

low throughput efficiency, limited life cycle capacity, and frequent need for maintenance, commercial pilots using lead-acid technology were deemed to offer too little benefit. Even with a limited scope, these initial installations revealed some valuable lessons about the installation, functionality, and long term viability of lead acid systems as well as providing the basis for the specifications for the storage equipment, power electronics, and communications systems of the forthcoming lithium ion systems.

### Subtask 1 – Select Candidate Sites

#### *San Carlos:*

The first installation site was a home in San Carlos, CA. The home was located in Pacific Gas and Electric Territory and the City of San Carlos was the Authority Having Jurisdiction (AHJ). During a preliminary site audit the home was confirmed to be a suitable location for the pilot based on two key factors: the customer had an existing PV system that could be easily integrated with energy storage equipment, the garage had seven lateral feet of wall space, and the electrical system in the house was simple enough that transferring critical loads to the firm PV system.

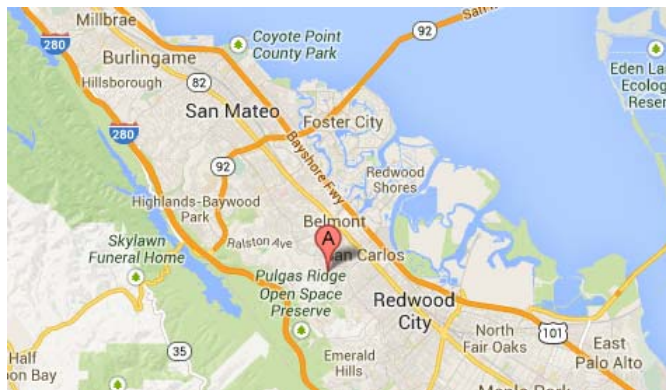


Figure 1 – Pilot 1 location



Figure 2 - Pilot 1 site





Figure 3 - Pilot 1 installation in process



Figure 4 - Pilot 1 battery enclosure

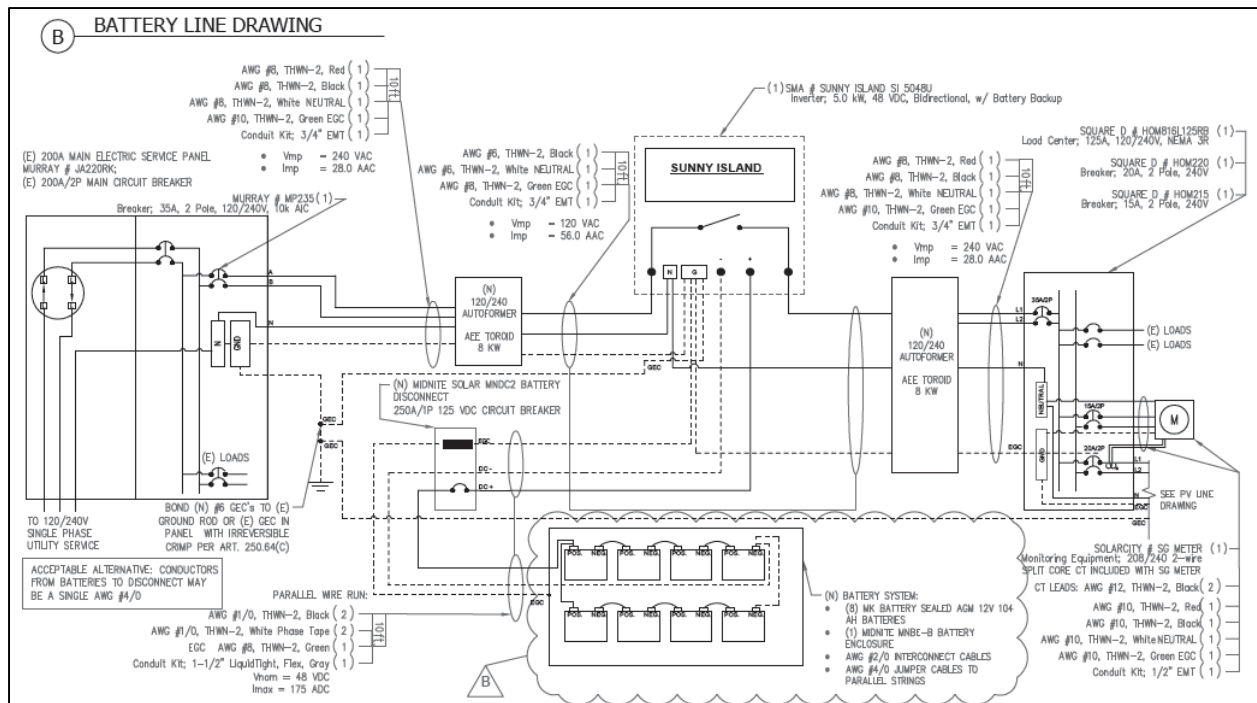


Figure 5 - Pilot 1 single line drawing

The installation process involved switching the existing Xantrex inverter with a Sunny Boy 3800 to enable optimum performance with the Sunny Island inverter/charger for the energy storage portion of the system. Due to the complexity of the system and mounting, the installation required several days.

### Oakland:

The second installation was in Oakland at the home of a SolarCity customer. This site was also in Pacific Gas and Electric territory. The City of Oakland was the AHJ and its planning department that not had seen a battery job of this kind. They were very interested in taking a close look at the equipment and installation process. The initial site audit revealed that although a fairly long wire run would be needed from the mail panel to the energy storage location, it would make a good pilot site.



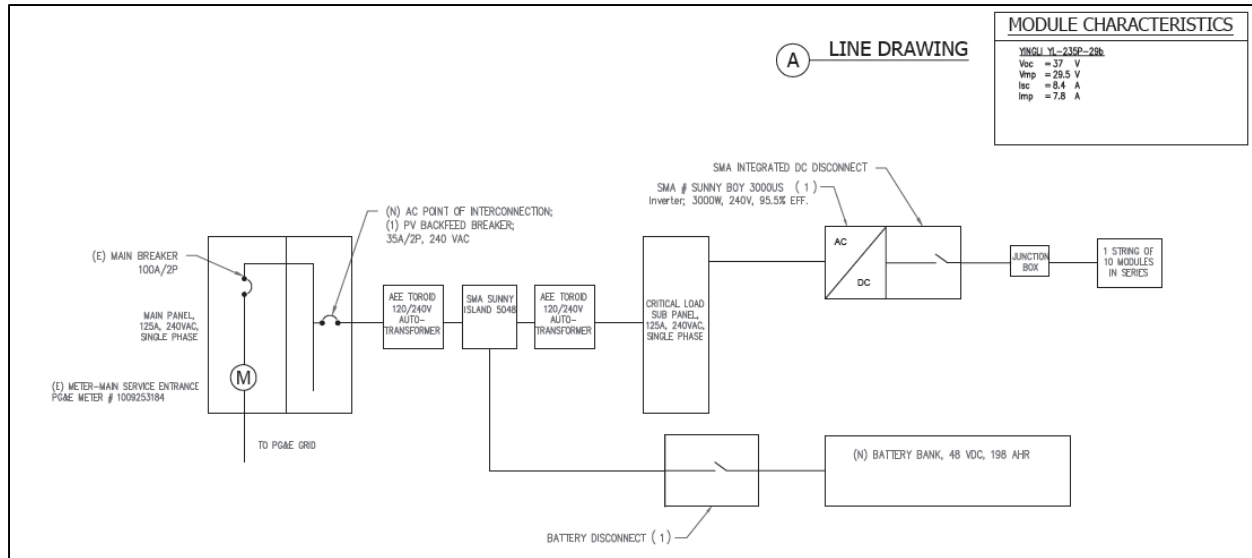
Figure 6 - Pilot 2 location



Figure 7 - Pilot 2 installation in progress



Figure 8 - Completed installation of pilot 2

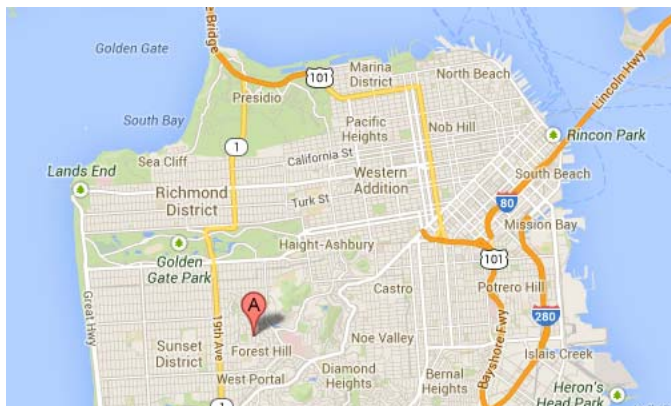


**Figure 9 - Pilot 2 single line drawing**

The installation process was simpler than pilot 1 because there was no need to replace the existing PV inverter. There was also ample wall space at this location which improved the ease of installation. There had been electrical upgrades when the PV system was installed.

### San Francisco

The third pilot site was located in San Francisco. This site was also located in Pacific Gas and Electric territory. The AHJ was the City of San Francisco which had a fairly easy process for dealing with permits for similar PV-paired energy storage systems. At the inspection stage two of the Senior Building Inspectors came out to the final inspection out of interest in the configuration and operation of the system. The audit at this site revealed a very easy installation with both the main electrical panel and the existing PV inverters in the garage where the energy storage system was to be located.



**Figure 10 – Pilot 3 location**



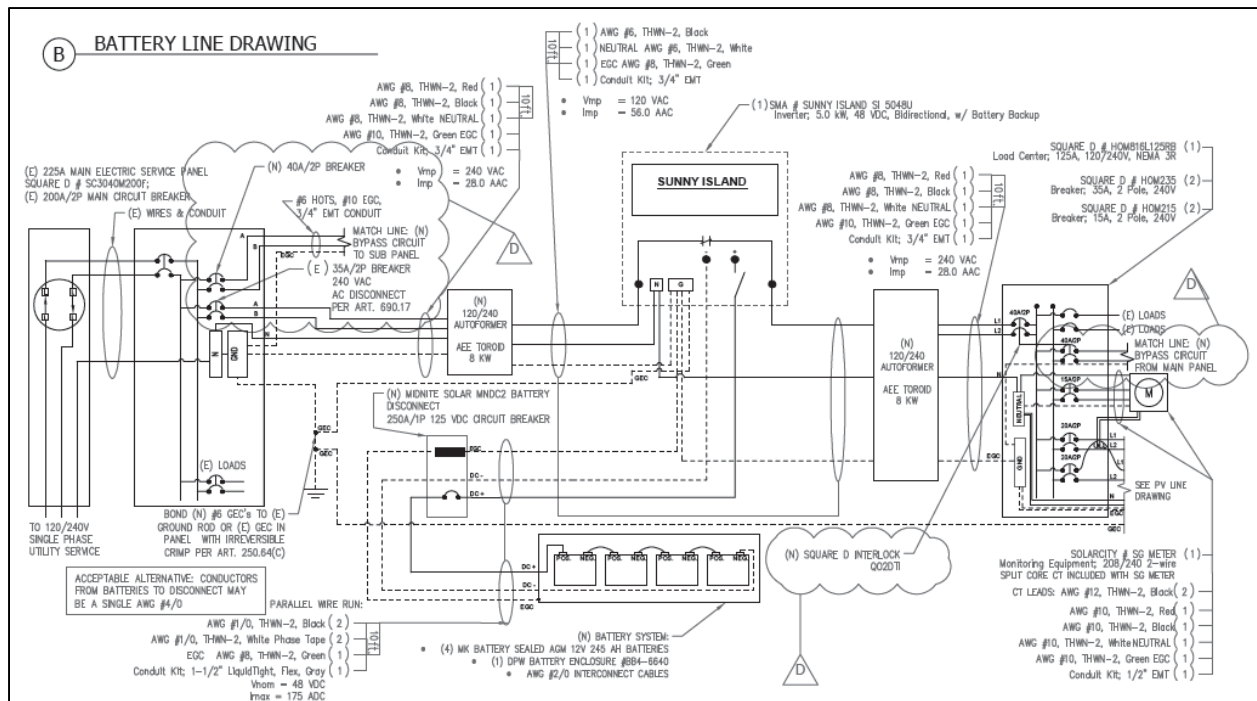


Figure 11 - Pilot 3 single line drawing



Figure 12 - Completed installation of pilot 3

The installation of pilot 3 was also complicated by the need to replace the existing Fronius inverter with an SMA Sunny Boy. There was also limited wall space and floor space to accommodate the installation so the placement of equipment and wiring was much more complex on this site than on the other two.

## Subtask 2 - Selection of Power Electronics

The selection of power electronics was a fairly straightforward process given the limited options currently on the market. The following vendors and models were considered:

Vendor	Model
SMA	5048U
Schneider	XW6048
OutBack	GVFX3648

The Schneider option was eliminated at the time due to a bug in the inverter firmware (that has since been fixed) that was causing voltage overshoot on the DC battery bus. The OutBack option was eliminated due to lack of support at the time for grid interactive programmable control.

The SMA Sunny Island 5048U was selected as the inverter/charger for the lead acid pilot systems based on scalability, efficiency, cost, availability technical support, and communications capability. Key concerns about the SMA product were:

1. 120 VAC Native Output – A 120 VAC output required the use of costly, bulky autotransformers to interface the Sunny Island with a 240 VAC PV inverter and native 120/240 VAC split-phase loads.
2. Indoor rated – The inverter is rated for indoor installation only, restricting options for pilot site installation
3. Heavy weight – The product weighs 139 lbs versus 121 lbs for the Schneider product and 62 lbs for the OutBack product. Excessive weight complicates transport and installation.
4. Limited integrated wire management and BOS – Unlike the Schneider and Outback options, the SMA product did not have off-the-shelf accessories for wiring and over current protection requirements.

### Subtask 3 – Battery Cells and Balance of System

Flooded lead acid and Valve-Regulated Lead Acid (VRLA) battery cells were considered for deployment. Flooded cells typically offer high energy capacity and cycle life per unit cost compared to VRLA cells<sup>4</sup>. A VRLA cell does not require watering, emits less hydrogen gas and is not at risk of spills. VRLA, Absorbed Glass Mat (AGM) type cells were chosen based on a lower expected lifetime cost, less mandatory maintenance, and design flexibility.

Additional Balance of System components included a utility AC disconnect, a battery DC disconnect, autoformers, and an enclosure for the batteries.

### Subtask 4 – Audit, Design, and Permitting

The systems were designed using a combination of information from existing photographs and information from the PV system installations and new information gathered at on-site audits. The key pieces of information needed to design these systems were: specifications of all existing PV system equipment, measurements and locations of desired mounting surface, electrical system specifications (main panel rating, brand, and location), and desired critical loads.

<sup>4</sup> Valve-Regulated Lead-Acid (VRLA): Gelled Electrolyte (Gel) and Absorbed Glass Mat(AGM) Batteries, <http://www.dekabatteries.com/assets/base/1927.pdf>

Once these data points were gathered, a designer created plan sets for each site showing a site plan, the PV system electrical diagram, the battery system electrical diagram, cutsheets showing specifications of all equipment being installed.

These drawings were delivered to the local AHJ and when possible, an electrician walked through the drawings with the plan checker or inspector. Acquiring permits was relatively straightforward for these installations and the permitting fees were between \$100 and \$300 with a wait time of 1-4 weeks.

### **Subtask 5 – System Installation and Inspection**

All three of the pilot sites were relatively similar in scope and complexity of installation. All three sites were in garages with existing PV systems to tie into. They were all designed to operate in both grid-interactive and back-up modes.

The steps of the installation included replacement of existing PV power electronics (when applicable, for compatibility with battery power electronics), the mounting and wiring of battery inverter/charger, mounting and wiring of autotransformers, installation of battery enclosure and batteries, and wiring of customer critical loads. The most time consuming element in all installations was the process of wiring the critical loads into the battery system. However, the additional cost of whole house backup is significant enough that it outweighs the added installation cost of rewiring critical loads circuits.

The interconnection applications to PG&E for these pilot installations were submitted as revisions to the existing PV systems. The applications were submitted with plans that detailed the systems being installed. These applications were all approved for interconnection and all sites received their official notice of “Permission to Operate” or PTO within approximately one month from the date of final inspection.

### **Lessons Learned**

These first pilot installations provided valuable information for the design of a lithium ion energy storage product including which features of a battery are most important and useful for a customer, and what is needed to integrate and operate with existing power electronics equipment.

Space constraints proved to be a significant concern with the initial pilot installations. Not only was the floor space required for the battery packs and enclosure an issue in many homes, but the wall space needed to install the power electronics and balance of system was a challenge for design and installation. Floor space was an issue for batteries on the floor of a garage, which is often the most suitable space for an energy storage system, not only because most residential customers have limited garage space, but because even when the systems could fit, there was a vehicle collision concern. Although there is typically ample room to mount this type of equipment outdoors, the cost and selection of outdoor rated equipment or enclosures is prohibitive.

For lithium ion systems, a system that can be wall mounted inside a garage as well as outdoors is necessary to make a storage product accessible to the majority of homes. It is also important that the system not require additional wall-mounted components such as discrete disconnects, wiring boxes and

autotransformers. Based on SolarCity's experience with typical customer wall space, the complete PV and battery system should be able to fit within a space of no more than four to five horizontal feet.

Another important lesson from the pilot installations was in the selection of power electronics. A native 120/240 VAC split-phase power output product would have simplified the interconnection process greatly. The power electronics with partial 120 VAC support required significant additional cost and installation space.

The final insight for the lead-acid pilot sites was the importance of compliance with the National Electric Code and importance of UL Listed equipment. Because so many energy storage systems have been off-grid in the past, many of the technologies that were part of these installations have not been up to code standards. As an example, the autotransformers that were required to convert 120 VAC to 240 VAC were UL Recognized but not UL Listed components. A field UL Listing was required to meet the requirements of a permitting authority. Field listings are time consuming and carry a heavy cost, severely impacting the economics of a project. To be commercially viable, a lithium ion system must incorporate both a battery and power electronics that have received a factory UL Listing. In SolarCity's experience, UL 1741<sup>5</sup> is the appropriate listing standard for battery inverters and UL 1973<sup>6</sup> is the appropriate listing for battery modules.

### **Lead-Acid Battery Technology for Commercial and Industrial Applications**

The project scope originally included deployment of three commercial and industrial pilot sites using lead acid technology. The pilots were not deemed beneficial and thus not deployed due to the following characteristics of lead-acid batteries.

**Limited Cycle Life** - A primary reason is the limited cycles over the system lifetime that lead acid products tend to be capable of achieving. Commercial sites are generally incentivized to operate batteries in a much more intensive way than residential sites. Demand reduction can cycle a battery multiple times per day and can require a deep discharge. The limited cycle life – 600 to 800 lifetime cycles at 80% depth-of-discharge – of lead acid systems makes the technology very limited in providing value to commercial sites. In order for a storage system to be commercially viable in this use case, a storage technology with lower cost per cycle is needed.

**Limited Warranty** - A typical lead acid storage product is warranted for between one and five years. A warranty of this term significantly limits the ability to finance the equipment which is a component of making energy storage economically feasible for commercial sites. Finance partners typically require a 10 year or longer warranty.

**Frequent Maintenance** - Another concern with lead acid storage at both residential and commercial locations is the maintenance required to keep the systems safe and functioning. Lead acid systems require servicing every six months to one year depending on whether the system is a flooded or sealed

---

<sup>5</sup> "Inverters, Converters, Controllers and Interconnection System Equipment for Use With Distributed Energy Resources", <http://ulstandardsinfonet.ul.com/scopes/1741.html>

<sup>6</sup> "Batteries for Use in Light Electric Rail (LER) Applications and Stationary Applications", <http://ulstandardsinfonet.ul.com/scopes/1973.html>



configuration<sup>7</sup>. Commercial customers tend to be very resistant to adding new technologies that require significant maintenance and the frequency of maintenance required makes a third party maintenance method financially unfeasible.

**Peak Power Limitations** - The primary use case for energy storage in the commercial market is demand reduction. Commercial utility rates are structured to incentivize reduction of peak demand. This means commercial storage systems need to be able discharge their energy fairly quickly (high power) in order to effectively shave these demand peaks. Due to high internal resistance, lead acid systems tend to have very low throughput efficiency at high power (loss of 25% or more at two hour, C/2 discharges)<sup>8</sup>. This significantly reduces the effective energy capacity of the system.

**Low Voltage Limitations** - Because individual cell maintenance is required, the National Electrical Code requires that these systems to operate at a nominal 48 VDC maximum<sup>9</sup>. This results in high amperage to output the same power. The high energy loss as well as increased equipment costs in high amperage, low voltage systems is damaging to the economics of demand reduction at current utility rates.

**Large Physical Size** - The final limitation of lead acid systems for commercial use is the large indoor physical space or costly enclosures required for these systems. A lead acid battery can only deliver 60 – 75 Wh/L versus an advanced lithium ion system's 250-730 Wh/L. The significant bulk of lead acid requires dedicated "battery rooms" at commercial scale or if indoor space is not available, requires complex outdoor rated enclosures for which the options are limited and the costs are extremely high.

## Conclusion

Lead-acid technology has been around for many years and is the technology of choice for backup systems for solar PV. However, there are many constraints with this technology that can be remedied by leveraging the huge improvements that both the automotive and consumer electronics markets have created in lithium ion technology, and in other novel battery chemistries. The pilot installations discussed in this section were integral to giving both Tesla and SolarCity the insights needed to develop a lithium ion battery pack and select system components that would create a cost effective and space constrained energy storage solution.

---

<sup>7</sup> LEAD-ANTIMONY, LEAD-CALCIUM, LEAD-SELENIUM, VRLA, NI-CD. WHAT'S IN A NAME?, [http://www.battcon.com/PapersFinal2009/ClarkPaper2009FINAL\\_12.pdf](http://www.battcon.com/PapersFinal2009/ClarkPaper2009FINAL_12.pdf)

<sup>8</sup> Effects of variability and rate on battery charge storage and lifespan, [http://dataspace.princeton.edu/jspui/bitstream/88435/dsp01x633f110k/1/Krieger\\_princeton\\_0181D\\_10531.pdf](http://dataspace.princeton.edu/jspui/bitstream/88435/dsp01x633f110k/1/Krieger_princeton_0181D_10531.pdf)

<sup>9</sup> National Electric Code, 2011, Article 690.71

## 2.2 Internet-Based Storage Control Platform

In task 2.2 SolarCity developed a control platform that enabled remote control of energy storage devices. The task involved developing the system architecture, implementing end-to-end communications infrastructure, and then implementing a suite of high level control algorithms based on the end use-case. Although the actual energy storage technology is a significant part of the challenge of functional and economic deployment of battery storage, the ability to control these systems is the key needed to unlock the value that these systems can provide to customers, utilities, and the grid as a whole.

### Introduction

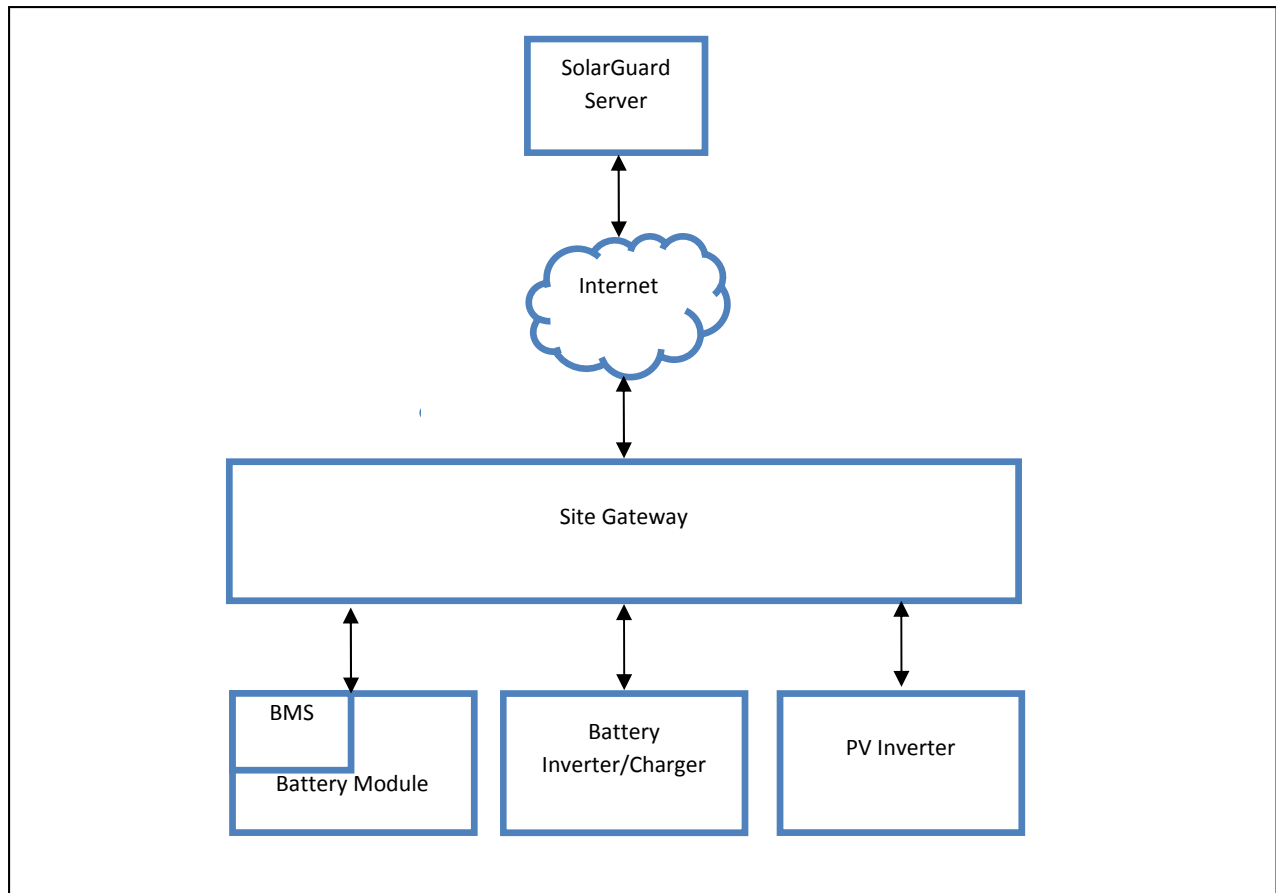
The communication platform for the SolarCity energy storage system builds upon the SolarCity SolarGuard PV monitoring platform. The existing SolarGuard platform provides real-time monitoring via a central server for SolarCity PV systems. The communication system adds modular, scalable support for inverter/charger and Battery Management System (BMS) protocols, onsite control intelligence, and bidirectional server control for a fleet of solar and storage systems.

During the design and preliminary implementation of the battery systems, SolarCity discovered a significant lack of industry accepted communication standards at each link within the system. Because of this, the initial implementation required use of a number of proprietary protocols. This section highlights standards could be adopted and where new standards are required. When discussing specific components, discussion is focused on commercially available Off-The-Shelf (COTS) components intended for generation and storage power densities less than 100 kilowatts.

### Communications System Architecture

Figure 1 shows the key components of the battery communications system. The components include:

1. Battery Module and Battery Management System (BMS)
2. Battery Inverter/Charger
3. PV Inverter
4. Site Gateway
5. SolarGuard Server



**Figure 13 - Communications system architecture**

### **Battery Module and BMS**

The BMS is typically packaged with the battery. It provides data acquisition and safety control for the battery system. The BMS provides charge and discharge limits to the battery inverter/charger.

### **Battery Inverter/Charger**

The battery inverter/charger controls charge and discharging of the battery. The inverter charger receives control commands from the site Gateway and limits from the BMS.

### **PV Inverter**

The PV inverter provides basic information on PV power generation.

### **Site Gateway**

The Gateway serves a central role in the communication systems. It performs the following functions:

1. Data collection from onsite components including the battery system, PV inverters, and load meters
2. High speed, secure interface to server for performance logging and central control over either IP based cellular or existing broadband connections

3. Local intelligence and control for times when a server connection is not available
4. Arbitration between BMS and inverter/charger in cases where the systems are not directly compatible

The Gateway supports a wide variety of industry standard hardware communication interfaces including RS232, RS485, CAN, ZigBee and Ethernet, providing support for a broad range of interfaces with system components. Similarly, the Gateway software is designed to be modular to support a plug and play system design. The software design supports a heterogeneous system architecture that self-configures on system startup. This enables a wide range of storage capacity or hardware components using the same Gateway solution.

### **SolarGuard Server**

The SolarGuard server provides centralized coordination and scheduling of a fleet of battery systems. The server also provides long-term logging and analysis of system performance.

### **Communication Protocols**

In the preliminary implementation of the battery control platform, it was quickly discovered that few open communications standards exist and none are widely adopted for any of the components within the system. Public standards reduce development, integration and interoperability testing time and cost. A single manufacturer can integrate multiple components using proprietary protocols and present a unified standards compliant interface, but this prevents quickly switching to new, best-of-breed subsystems as they become available.

The following sections discuss the individual interfaces between each component and suggested areas for standards development.

### **Gateway to PV Inverter/Charge Controller**

COTS PV inverters have communications interfaces requiring a large variety of physical interface standards including RS232, RS485, Bluetooth, and CAN. Supporting a large variety of physical interfaces increases the complexity and cost of the Gateway. Additionally, wired interfaces add additional cost as it is not always advantageous or possible to mount the Gateway in close proximity to other components.

COTS inverters also use a variety of proprietary communications protocols. Even in cases where manufacturers have chosen public standards such as Modbus<sup>10</sup>, there is significant variability the implementation.

*Suggested Protocols: SunSpec Inverter Monitoring and Control<sup>11</sup>*

---

<sup>10</sup> [www.modbus.org](http://www.modbus.org)

<sup>11</sup> [www.sunspec.org](http://www.sunspec.org)

### **Inverter/Charger to Battery Management System**

Battery inverter/chargers and BMS systems both use a variety of physical layer and higher layer protocols. Even in cases where both vendors use a similar protocol, CAN<sup>12</sup> for example, there are many options for incompatibility in the implementation.

*Suggested Protocols: CANopen CiA 418/419<sup>13</sup>*

### **Gateway to Inverter/Charger**

The Gateway to inverter/charger interface should adhere to the same specification as the Gateway to PV Inverter interface. Current COTS inverter/chargers implement a variety of proprietary protocols. Additionally, published support for grid-interactive control is limited.

*Suggested Protocols: SunSpec Storage & Charge Control*

### **Gateway to Server / Server-to-Server**

The Gateway to Server link is often managed by a single entity, the system owner and operator, thus may be customized to meet the owner's specific requirements. However, the operator may pass control of the battery system, or a fleet of systems to a third party (e.g. the balancing authority), and thus a standardized interface for third party control is required.

There are well accepted standards for security, authentication, and encryption over public IP networks. The SolarCity battery system implementation is designed to adhere to the security policy described in the OpenADR specification<sup>14</sup>.

*Suggested Protocols: OpenADR 2.0*

### **Conclusion**

Building off of the existing SolarGuard platform, SolarCity was able to add modular and scalable support that enabled the power electronics and battery management systems to both relay data, and take commands from a SolarCity server. Much can be done to simplify and unify industry protocols and specifications by which each piece of a battery system communicates. As discussed in the recommendations above, SolarCity found that open standards that allow interchange of different system components from different manufacturers will be critical to advancing battery systems and especially in reducing system costs.

---

<sup>12</sup> ISO 11898

<sup>13</sup> [www.can-cia.org](http://www.can-cia.org)

<sup>14</sup> [www.openadr.org](http://www.openadr.org)

### **3 Develop and Deploy a Grid-Interactive Li-ion Battery System**

Task 3 encompassed the core technology development and demonstration goal of the project. Tesla adapted the advanced battery technology used to produce vehicles and developed a stationary li-ion battery pack that could be used in both residential and commercial applications. In the second phase of this task, SolarCity completed the process of locating, designing, permitting, building, and interconnecting the li-ion systems. Although there were significant challenges along the way, the outcome was successful applications of li-ion battery systems for both residential and commercial customers.

### 3.1 Integrate the Tesla Motors lithium-ion, high-voltage battery pack with a grid Interactive inverter, charger, and PV System

#### Project Goals

##### *Objective*

As identified in the grant application, the goal for Tesla's share of the project was to "significantly advance a battery storage technology that has been refined in the competitive automotive industry". This was to have been shown through the demonstration of an integrated PV and storage product and to demonstrate an advanced lithium ion (Li-ion) battery storage platform. These goals have been achieved, on time and within budget as proposed.

##### *Summary of Developments*

In order to adapt a li-ion battery which was developed for the automotive market and to demonstrate the product, several key developments were required, as follows.

##### *Hardware*

The automotive battery enclosure is designed to mate to a vehicle in a vehicle production environment, not to be friendly to residential site conditions and installation personnel. The main interface, the battery enclosure, therefore had to be redesigned for this new environment and use case.

The first step in this process was to remove the essential battery components from the automotive enclosure, which at the time was from one of Tesla's battery development programs with an OEM partner.



The original vehicle battery enclosure.



The vehicle enclosure has been removed, and the modules have been reconfigured electrically to operate at 48V (shown here with the residential inverter/charger, the yellow box). The “cakestand” is only to protect the modules in testing; subsequently, the formal design of the new enclosure was further refined:



New enclosure, with modules in upright orientation.

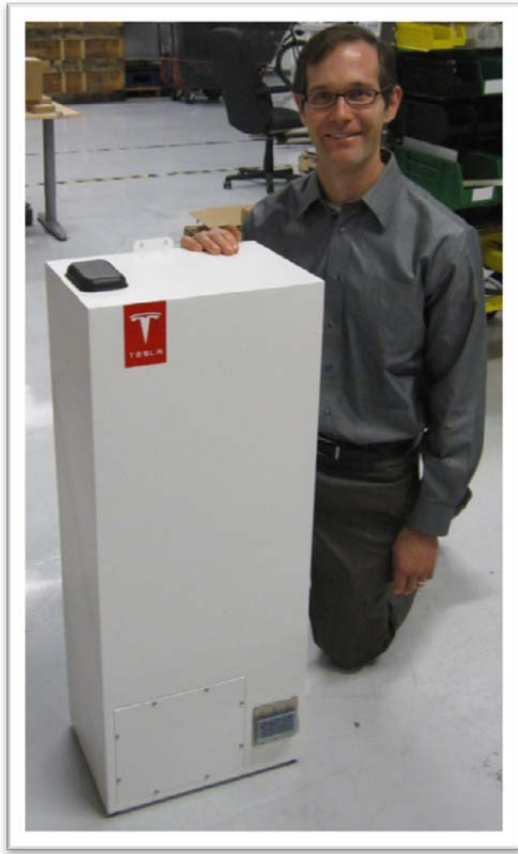




However, this initial enclosure design was built around a custom module design which used one of Tesla's battery development programs for Daimler. At this time, this module design represented the state of the art for Tesla. As newer module design was available due to the development progression of the Model S vehicle, this design was modified to incorporate the Model S modules, which would be made with higher quality and lower cost.



The final enclosure design under the CSI program was the following:



### *Electrical interoperability*

To integrate with existing equipment and meet the power needs of a residential setting the DC voltage of the residential pack had to be reduced from the automotive pack. The automotive battery pack operates at a nominal voltage of 350V DC and its electrical architecture is designed to meet the power and energy requirements of driving. However, the power and energy requirements of home appliances are lower and this dictates adjustment to the pack architecture. Furthermore, the existing bi-directional inverters on the market for low power applications like this one were designed for lead acid batteries with a nominal operating voltage of 48V DC. The residential pack is able to use the same basic architecture at the battery module level as the vehicle, with modification. The residential pack uses fewer modules in series and has one module with fewer sub-units, bricks, to lead to the desired nominal voltage of 56.25V DC. This voltage is still higher than lead acid but it is still compatible with existing inverters. While this particular change was not dramatic, Tesla nonetheless still had to design new parts, create new manufacturing process and test for safety.

### *Software/Communication*

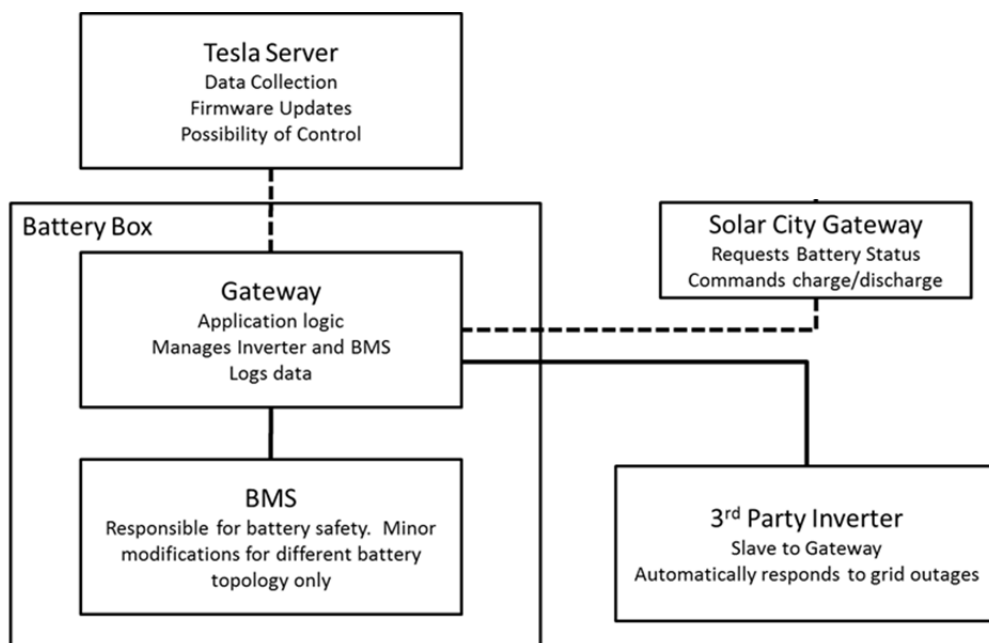
The automotive Battery Management System (BMS) which Tesla had at the start of the CSI program was designed to operate in vehicle conditions and through standard automotive formats, specifically and exclusively CAN. No capability existed for the battery to communicate with an off the shelf 48VDC residential inverter/charger, or for the battery to be integrated into

a network such as SolarGuard. As such, an additional layer of communication and control had to be incorporated into the product, along with modifications to the BMS itself. Inverter communications with the Schneider Xantrex XW were developed over Xanbus a 29-bit CAN interface and higher level protocol. After development this communication path was tested and idiosyncrasies of the Xantrex XW Tesla battery pair were smoothed out. In addition, Tesla has developed a wireless communication interface using XBee radios and the ZigBee higher level protocol. This wireless interface allows a third party, like Solar City, to issue charge and discharge commands while it monitors the solar PV supply and load at an installation site.

The Tesla battery will obey the operator commands as long as it is not unsafe the battery to operate at the requested value. Tesla also provides basic feedback to the third party about the battery's state of charge and other vital pieces of information, to help them make better dispatch decisions.

In addition, the Tesla Gateway provides interface between the inverter the BMS and the third party operator and provides additionally functionality that is unique to the residential application, and separate from automotive operation. The Gateway controls fault clearing and re-try behavior to allow the inverter/battery system to operate more reliably. The gateway can also operate the battery independently in back-up mode. For example, in the case of a power outage where the battery – solar PV pair revert to back-up mode and the solar PV is providing more power than the battery and the loads can accept the gateway will issue a command to the Xantrex inverter which increases the AC frequency out of spec so that the solar PV inverter will turn off and the battery can rebalance the supply and demand of the system.

Developing all of this functionality required close collaboration between Tesla's stationary storage engineers and the vehicle BMS engineers, as well as with SolarCity.



### Technical Progress since CSI RD&D

Since the end of the development under the CSI program and thanks to the progress made under that program to establish the technology platform, Tesla has continued to develop the product, and deploy systems into the market. Tesla's current Stationary Storage Residential Unit is a 10kWh AC battery with a maximum discharge power of 5kW. It is compatible with a Schneider Xantrex XW 6048 bidirectional inverter and as was intended from the demonstration program, the unit as developed can be used for load shifting, back-up power during off-grid events, and integration with PV to provide firming support and load shaping.



Specifically, there were significant improvements in mechanical design, manufacturing process, communications and controls. Design adjustments have been based on the learnings from the field under the CSI program, as well as Tesla's own internal testing. The production rate has increased to 10 packs per week and is in the process of transitioning to full scale production. Additionally, field installations have begun with field listing to UL 1973 UL testing is underway to become a UL certified product by November 2013. A communication interface for 3<sup>rd</sup> parties, currently used by Solar City has also been developed and tested and continue to improve inverter controls. Importantly, Tesla has also created a viable roadmap to the very aggressive price target of <\$500/kWh established under the CSI program.

### Design for manufacturability

The design and manufacturing teams at Tesla are able to closely collaborate because of their small size and close proximity, and this initial relationship was established through the CSI program. This collaboration allows manufacturing to give feedback to the design team which can quickly be rolled into the next design. Small adjustments like changing antenna length and rearranging components in the wiring compartment have been made seamlessly. As a transition to full scale production is planned the design team will continue to be involved in making changes to the product to facilitate efficient production and meet cost targets.

### Manufacturing Process Improvements

One of the most visible changes in the last year is the development of a New Product Introduction (NPI) manufacturing line for the Residential Stationary Storage Unit. With the involvement of Manufacturing Engineering, Quality Engineering and more man-power on the line itself Tesla has created the processes and built the tools required to build at a rate of 10 packs per week. This includes creating complete work stations with part buckets, manufacturing instructions and tools laid out for each part of the build process. NPI has also set up and continues to develop End of Line (EOL) testing to quickly



verify that each product is fully functional. Another addition is the pack build carts, pictured above, which are now seen all over the build area and convey a pack from start to finish. NPI has also begun outsourcing certain parts, like wire harnesses, and are planning to automate certain processes in house. In November the production of these packs will move to the Tesla factory in Fremont, CA. This move will be a major step in the ongoing process of transferring from NPI to full scale production.

### *Field installations*

Solar City is well underway installing Tesla Residential Stationary Storage Units in homes in California. Many customers already have Solar PV provided by Solar City and are eager to try the new storage technology. Some customers are also Tesla Model S owners enthusiastic about all Tesla products. About 50 batteries are currently installed in homes and many of those (about 40) have been UL field listed to UL Standard 1973. UL listing is a way to verify the safety of the product and is required before it can be interconnected. Tesla is delivering 10 packs per week to Solar City, who is able to quickly turn around and install these units, and as a result field deployments will steadily grow in the coming months. Tesla has already learned some important lessons from initial field units and having a fleet with real performance data will help to continue to develop and improve the product in meaningful and practical ways.



## **Appendix – Product Specifications**

### *Electrical*

#### System Level

Continuous power	5	kW AC
AC Capacity @ 5kW (92.5% Inverter Eff)	10	kWh AC

#### Battery Only

Voltage, nominal	56.25	V DC
$V_{min}$	40.5	V DC
$V_{max}$	63	V DC
Continuous current	120	A DC
Peak discharge current (15 sec)	250	A DC

### *Ambient Operating Temperatures*

Continuous Operation and Charging

Low Operating Temp	10 °C
High Operating Temp	40 °C
De-rated operation above 40 °C	
Discharging	
Low Operating Temp	-20 °C
High Operating Temp	40 °C

<i>Efficiency</i>	C/2	C/4
DC (battery only), roundtrip	96%	98%
AC (including inverter), roundtrip	81%	85%

*Backup / off-grid capability*

- Automatic backup / off-grid functionality
- Provides 1 week of backup power for an energy star refrigerator
- Potentially unlimited backup capability if used with solar

*Mechanical and Mounting*

- Height x Width x Depth: 1167mm x 484mm x 417 mm
- Weight: 168 kg
- Wall Mount
- Indoor enclosure NEMA 1 rated

*Regulatory*

- UL listing expected to UL 1973

## 3.2 Deploy integrated FirmPV / high-voltage storage systems

### Introduction

Task 3.2 focused on field implementation by SolarCity of the technology developed by Tesla in task 3.1. The experience SolarCity gained during the lead acid pilots discussed in task 2.1 informed many of the design and deployment decisions for the li-ion projects. Although there were some hurdles with permitting, inspection, and interconnection, all six sites were successfully deployed and are providing value in both the residential and commercial use cases.

Although a significant portion of system cost is driven by the cost of the battery pack, the costs of balance of system and installation are significant enough to make or break the economics of adoption. This task was important in helping to drive down these costs as well as to test the field performance of the technology.

### Subtask 1 – Select Candidate Sites

#### *Residential - San Francisco and Woodside*

Two of the residential li-ion pilot sites were located in San Francisco and one was located in Woodside. Like the lead acid pilots, these sites were located in Pacific Gas and Electric territory. The AHJs were the City of San Francisco with whom SolarCity had worked on one of the previous pilots, and the city of Woodside, which simplified the permitting process. One site was located at a home that had a relatively small electrical load so the battery system was designed to provide backup to the entire home, while at the other pilot sites only selected loads were connected to the battery system.

One of the primary challenges with the lead acid pilot installations was the amount of physical space required for the battery, inverter/charger, and autotransformers, in addition to the equipment and space needed for the PV installation. The design of the li-ion system focused on a product that could be relatively compact and wall mounted. This factor and the elimination of the need for autotransformers made selecting an installation surface much easier. Locating the systems near both the main electrical panel and the critical loads was the biggest constraint but both residential systems were successfully mounted on the garage walls of the two sites.



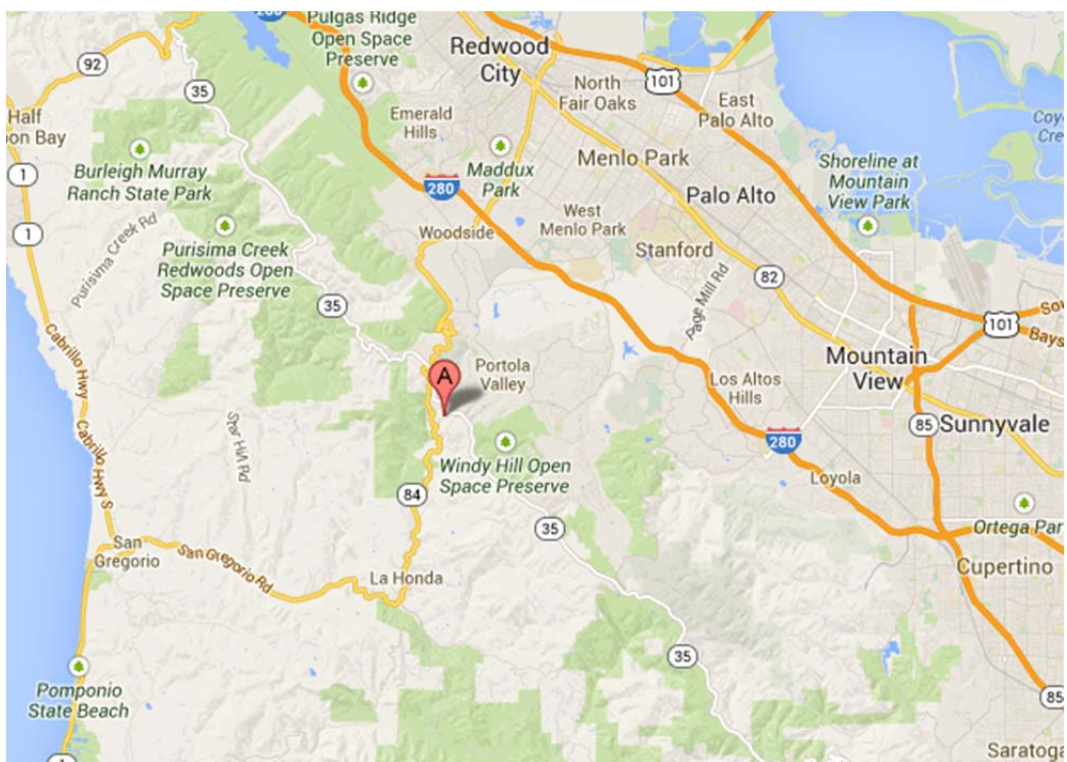
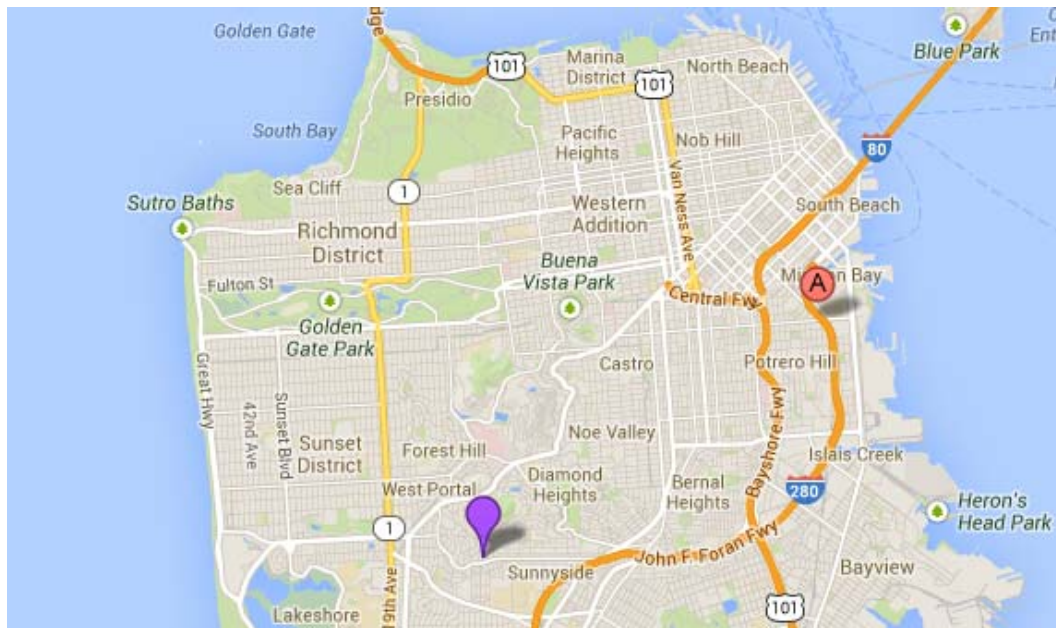


Figure 14 - Location of Residential Li-ion Pilots





Figure 15 - Residential Li-ion Pilot Installation



Figure 16 - Residential Li-ion Installation

The battery power electronics selected for the installations enabled more flexibility in integrating with the existing PV systems on site. Neither existing PV inverter needed to be replaced. In comparison to the lead acid installations, the physical space required and the complexity of the installations were greatly reduced.

### *Commercial – Northern and Central California*

The primary goal for the commercial pilots was to demonstrate that li-ion energy storage systems coupled with PV could effectively reduce peak demand. Although PV coupled storage can provide many benefits, peak demand reduction is the primary use case currently incentivized by utility tariffs. Providing backup power for certain loads can provide value to commercial customers but was not included as a feature of these systems. Additionally, participation in ancillary services markets which provides services to the grid is a key future use of the technologies being developed for this pilot.

The two commercial pilot sites were selected from SolarCity's existing PV customer base. The key criteria for selection were sites with significant demand charge rates and therefore good candidates to test the capability of storage to provide demand reduction. Both sites were located at retail stores with load of between 300 and 600 kW. The energy storage devices on these sites were paired with existing PV systems of 350-500 kW.

Both sites were located in California in Pacific Gas and Electric territory, which has comparatively high peak demand charge rates for commercial sites. One site was on the E-19S tariff which is a Time of Use (TOU) tariff and has a peak demand charge of \$14-\$16 per kW. The other site was on the A-10SX tariff which is a non-TOU rate and has a peak demand charge of \$10-\$14.



**Figure 17 – Northern California Li-Ion Pilot Installation**



Figure 18 – Central California Li-ion Pilot Installation

### *Commercial – Pennsylvania Location*

The other commercial pilot was installed at a commercial location in Pennsylvania. The site was located in the Pennsylvania Power and Light (PPL) electric utility. This site was similarly geared at peak demand reduction and designed to be operated in conjunction with a PV system on the same site and was installed in collaboration with PJM, the Regional Transmission Organization that coordinates the movement of wholesale electricity to parts of 13 states. PJM's Alternative Technology Resource Pilots program.



Figure 20 - Pennsylvania Li-ion Pilot Installation

## **Subtask 2 – Select Power Electronics**

The selection of power electronics for both residential and commercial pilots was informed both by the lessons learned during the lead acid pilot phase (Task 2.1). Since advanced energy storage technology is very much still under development, the availability of battery-enabled power electronics is limited. The primary manufacturers are Schneider Electric with the XW model, SMA with the Sunny Island (used in the lead acid pilot installations), and Outback Power's line of grid-tied products.

For residential installations, the key selection criteria were a native 120/240 VAC split phase output, software programmability for integration with the Li-ion Battery Management System, and device cost. The SMA Sunny Island was chosen for the first residential Li-ion installation due to familiarity from the lead-acid pilots. For the second two installations, the Schneider XW inverter was chosen due to native 120/240 split phase output.

For commercial installations, a three-phase, 208 VAC output for integration with the building electrical system was the most challenging constraint. For this reason, the SMA Sunny Island was chosen for ease of three-phase support.

In general, power electronics for integrated PV and energy storage systems are one of the more limiting factors in further development. Investment in further development of this crucial system component is a key lever in accelerating the growth of energy storage technologies and deployments.

## **Subtask 3 – Audit, Design, and Permit Systems**

With selection of sites and equipment completed, the next step was to audit each site to determine specific site constraints that would determine the system design. For the residential systems the concerns were similar to the lead acid pilot phase. Although the li-ion systems were significantly more compact and could be fully wall-mounted, it was still necessary to select a mounting location with at least 48" of horizontal wall space. Selection of critical backup loads was another key concern.

The first pilot had small enough loads, measured and verified for code compliance purposes, that it was possible to back up the entire home. Rather than having to pull loads from the main panel or from indoor subpanels, the home, the PV and the battery system were all relatively easily integrated. All equipment was located in the garage.

The second residential pilot site had plenty of internal mounting space in the garage. However, the electrical system was more complicated, including a load center in the garage and one inside the home on an upstairs floor. Although the whole home could not be backed up, a load monitoring device was placed on the subpanel to verify that the complete subpanel could be supplied by the storage system. After a month of data collection, the subpanel was verified to have a low enough draw that it could be wholly wired in to the energy storage system.

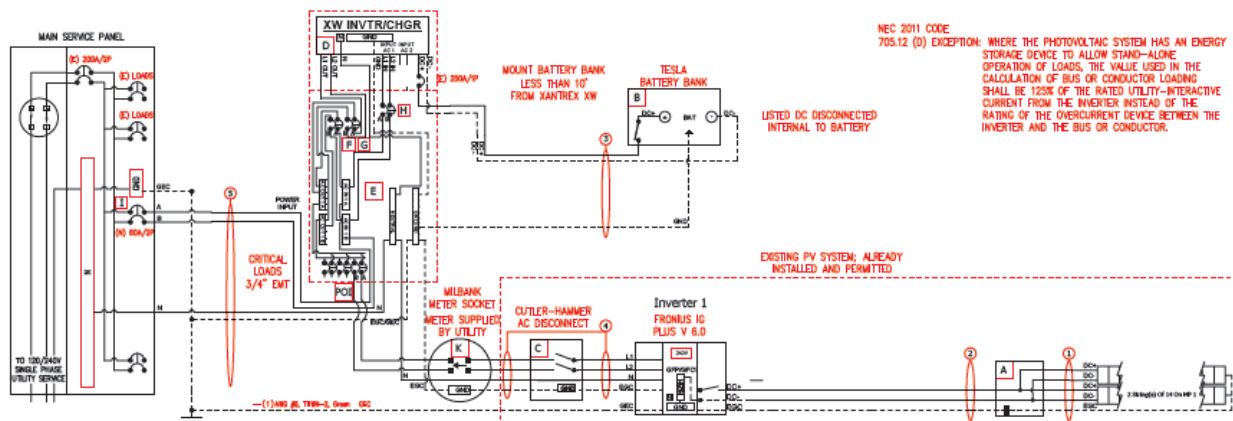


Figure 21 – Residential Pilot 2 System Line Drawing

For commercial sites, the key factor for system location and design was physical mounting space for the equipment. The systems needed to be indoors, wall-mounted, and in a location where bollards could be installed to prevent collisions with heavy equipment. The PV systems were not directly connected to the SMA inverter/chargers in these installations. The storage system was connected into an existing subpanel nearby. The primary concern for these sites was to provide peak load reduction, so rather than needed to connect to specific loads, it was only necessary to find a load center that was able to handle the back-feed of the storage system.

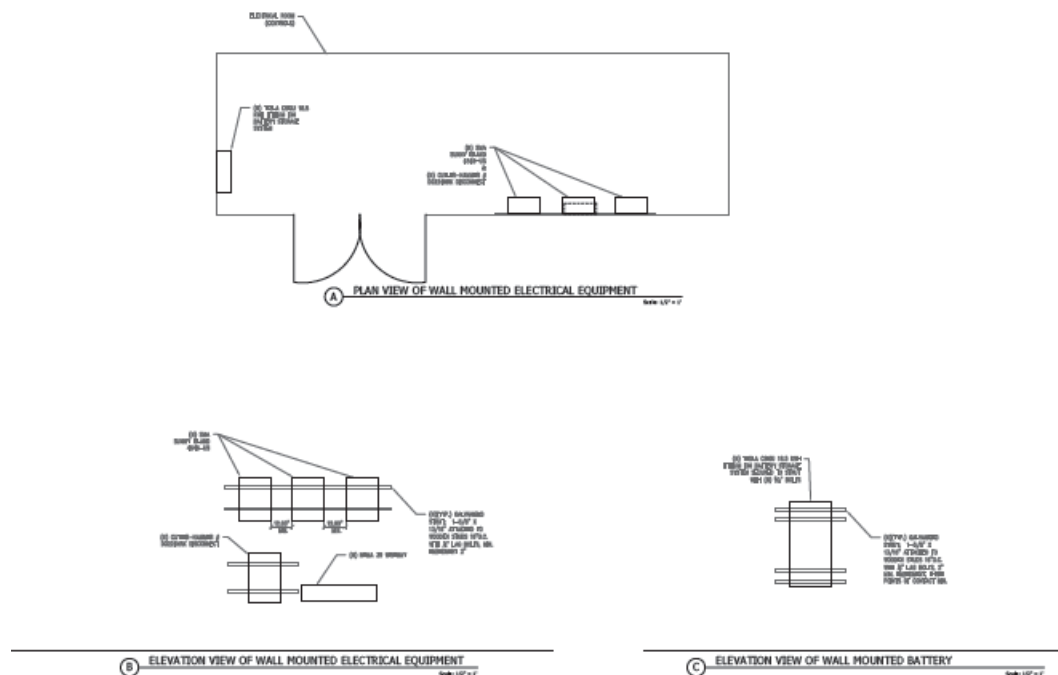


Figure 22 - Positioning of Commercial Li-ion Pilots

After gathering data from existing photo of the PV installations and on-site audits, the design process for both residential and commercial systems was relatively simple. The selection of equipment, location of

equipment, and integration with on-site electrical system completed, formal plans were created and submitted to the AHJ's for each site. Depending on local requirements these plan sets included site plans, wiring diagrams, load calculations, equipment specification sheets, and structural calculations.

#### **Subtask 4 – System Installation and Inspection**

The residential installations took a single electrician around four work days to install. This included significant process development and documentation. The installation process began with fixing a plywood board to the wall studs with lag bolts. The mounting hardware for the battery, inverter/charger and critical loads panel were mounted to this surface. Due to its weight, the battery was mounted directly onto two studs. The inverter/charger and integrated critical load panel were lighter and did not have to be mounted directly into a stud. The battery is then wired to the charge controller and the inverter is wired to the charge controller as well.



**Figure 193 - Mounting the Residential Li-ion System**

Once the major equipment was mounted, the PV output then had to be intercepted and diverted into the critical load panel for charging the battery and powering critical loads in an off-grid scenario. One other requirement from utilities was to add a Net Generation Output Meter (NGOM) onto the PV circuit to measure PV generation and ensure that Net Energy Metering (NEM) export credit was not given to battery discharge.



The next step was then to wire in the critical loads. For the two residential pilots, the fact that it was possible to back up an entire panel made the integration of critical loads much easier than it would have otherwise been. In this case, a 60A breaker was installed in the critical load panel and the load center was wired into this breaker.

The commercial installations required a scheduled shutdown at each site. Given that they are both retail locations, shutting down power to the whole facility can be difficult and costly. The installation itself was very similar to the residential systems with the exception of wiring critical loads or a PV circuit directly into the system.



**Figure 24 - Commercial Li-ion Installation Underway**

A series three of SMA Sunny Island inverter/chargers, a charge controller, and the battery were all mounted on steel strut. The equipment was then wired together in a gutter box and connected to a subpanel nearby. The utility also required installation of a NGOM for NEM administration which added significant cost of approximately \$3000.

The sites then went through a field UL listing inspection to verify that the Tesla equipment met all UL requirements. These visits were conducted by TUV Rheinland and were completed before the final inspection by the local AHJ. These inspections passed as did the local inspections with City of San Francisco, City of Stockton, and City of Mountain View. The inspection documents were then submitted to PG&E for all sites and after system inspection and installation of meters, these sites were given permission to operate (PTO).

### Subtask 5 – Test Completed Systems

Once PTO was received, each installation was then tested for performance. The equipment was tested to ensure that it was functional and up to the expected specifications. More importantly, the use cases were tested for backup purposes and load shifting on the residential side and for peak load reduction on the commercial side.

Market Type	Nameplate	Capacity	Efficiency	Power Electronics	Battery Type
Residential	5kW AC	10kWh	80%	SMA Sunny Island	Li-Ion
Commercial	9kW AC	18kWh	78%	Xantrex XW	Li-Ion

### Conclusion

The key goal of the deployment in this phase was to prove the field viability of the products that Tesla developed in Task 3.1. This task demonstrated several benefits of li-ion technology over the more traditional lead acid equipment discussed in section 2.1. The newly developed energy storage systems from Tesla enormously improved ease of installation due to decreased equipment size and weight. The ability of li-ion systems to perform over a long timescale and the lack of required maintenance is a key enabler of commercialization of energy storage. There were still several key challenges to deployment including inconsistent interconnection policy and limitation of power electronics options. However, the deployment in task 3.2 was largely successful and created a product and process that can be replicated at scale.



## 4 Monetization of the Value of PV-Paired Energy Storage

This section of the grant project focused on the economic value in battery storage paired with PV systems. Once the technology was developed and deployed, the key remaining component is the use cases and regulatory environment needed for the benefits of storage to be accessible. New technologies like PV paired with grid interactive storage have the potential to provide substantial cost savings for utilities, ratepayers and customers and reduce carbon emissions to a far greater degree than either PV or storage could achieve on its own, while also helping ease the strain on aging utility infrastructure. The benefits of this pairing include increasing the value of solar energy by enabling it to be effectively dispatched to times when it is more valuable to the system, reducing the costs of integrating solar energy into the distribution system and avoiding upgrade costs that might otherwise be required. The key to unlocking these benefits is overcoming the barriers to adoption, including upfront costs.

Remote, fully bidirectional control of distributed storage assets that can stabilize or ‘firm’ otherwise intermittent renewable resources will enable both site-specific peak demand reduction benefits and system-wide grid network benefits that will reduce costs and carbon emissions and improve grid reliability. As the penetration of PV systems increases, firming local solar electricity production can diminish the need to operate voltage regulating equipment in the distribution system (and thereby increase the lifetime of these devices). Distributed storage can also increase the stability of the grid from behind the meter and thus reduce the need for ancillary services from the California Independent System Operator (CAISO). It can also render additional capacity in the real time economic dispatch stack unnecessary. These measures hold significant promise for reducing the grid integration costs of distributed PV, and could also decrease grid-level emissions by moving the supply side to a more efficient operating state. Distributed, firming PV presents a path to secure, long-term and low cost carbon abatement, which can help the state meet its renewable energy and AB 32 greenhouse emissions reduction goals.

## 4.1 Advanced Energy Storage Market Survey

### Introduction

Significant numbers of SolarCity customers are willing to purchase battery storage products as a way to provide uninterrupted energy needs, to protect themselves against power outages and to reduce environmental impacts of energy use.

To assess interest in customer-sited battery storage systems, SolarCity surveyed 1,000 of its customers. 207 people responded. The results indicate that a large majority of these customers have an interest in battery storage (85 percent), are interested in purchasing (79 percent) a storage system and have a willingness to spend more for an energy storage or backup generation product (56 percent) that holds benefits for the environment.

As a strong indicator of the value of public-private collaborations on solar and storage technology, 85 percent of survey respondents said their interest in battery storage improved, knowing that rebates would be available through a state or municipal agency.

The unpredictability of traditional energy delivery was a major factor in customer interest in a battery storage system. Nearly all survey respondents, 96 percent of the total, said they had lost power at least once within the last five years, and 71 percent said power outages had occurred three times or more within five years.

Power losses raise a spectrum of concerns for consumers, with the largest numbers of respondents saying they worry about food spoilage in refrigerators, adequate heat in their homes and the overall comfort and convenience of everyday living.

Those concerns help explain why more than three quarters of all respondents, 77 percent, said they would be motivated to buy a storage system, knowing that it could help provide basic energy needs in times of energy outages.

## Methodology

A battery storage survey was conducted to probe potential customers of SolarCity battery storage products (BSPs) to gain deeper insight into the customer understanding, mindset, and desires for this for this product category. Current SolarCity residential Solar Photovoltaic and residential Energy Efficiency Upgrade customers were targeted for their opinions. The survey was emailed the first week of November 2012 to ~1,000 residential customers that had not received SolarCity communications in the previous 30 days. As an incentive, participants who completed the survey were given a \$10 gift card.

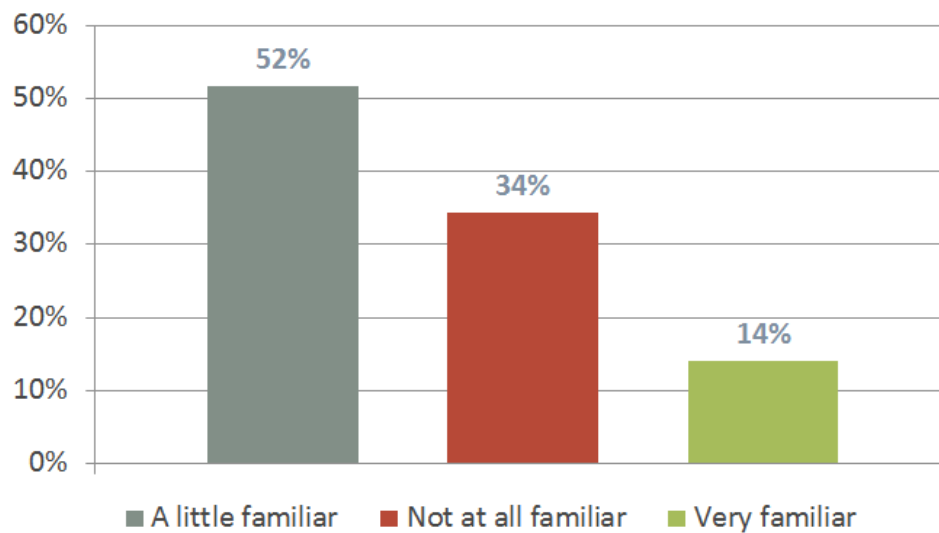
Survey Responses by	
State	Total
California	68
Connecticut	20
Hawaii	22
Maryland	32
Massachusetts	33
New York	32
<b>Grand Total</b>	<b>207</b>

## Survey Results

### Familiarity with Battery Storage

Two thirds of respondents said they were “familiar” with battery storage, with more than half of them (52 percent) saying they were “a little familiar” and 14 percent saying they were “very familiar.”

The remaining third, 34 percent, said they were “not at all familiar” with battery storage.



N= 207

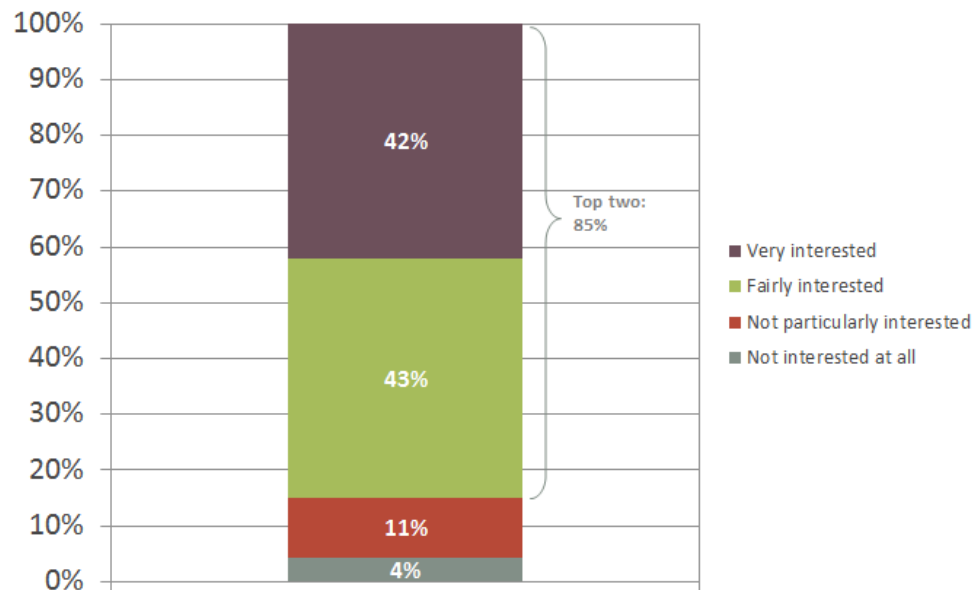
Q2 in Solar Survey,  
Q2 in EE Survey

**Figure 1. Responses to the question regarding familiarity with battery storage**

### Initial Interest

An overwhelming number of respondents, 85 percent, expressed interest in the idea of battery storage, evenly divided between those who said they were “very” interested (42 percent) and those who said they were “fairly” interested (43 percent).

Only 4 percent said they had no interest at all.



N= 207

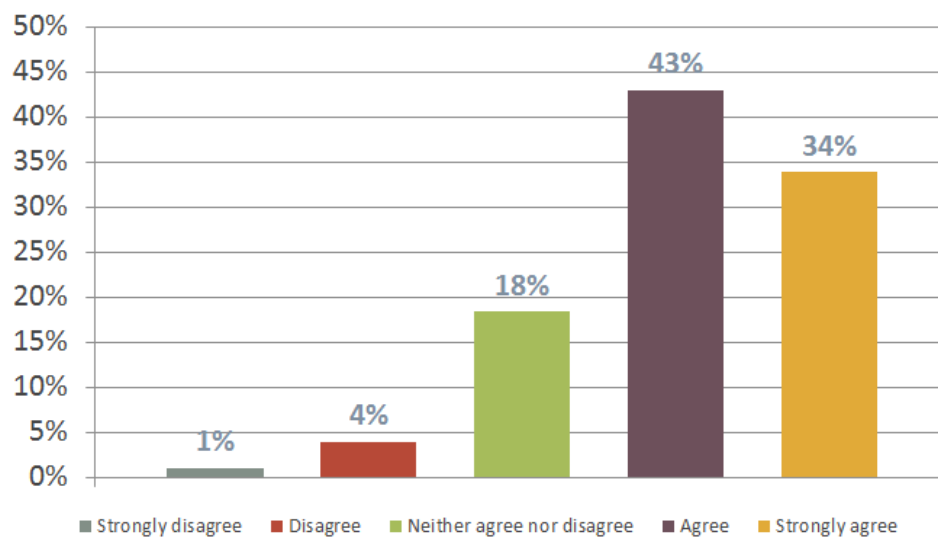
Q3 in Solar Survey,  
Q3 in EE Survey

**Figure 2. Gauging interest in battery storage**

### *Solving an Immediate Problem*

More than three quarters of respondents, 77 percent, said knowledge that a battery backup system can provide basic home energy needs for a few days if power goes out would motivate them to get such a system.

Another 18 percent were open to the possibility getting one while only 5 percent of respondents ruled out a purchase.



N= 207

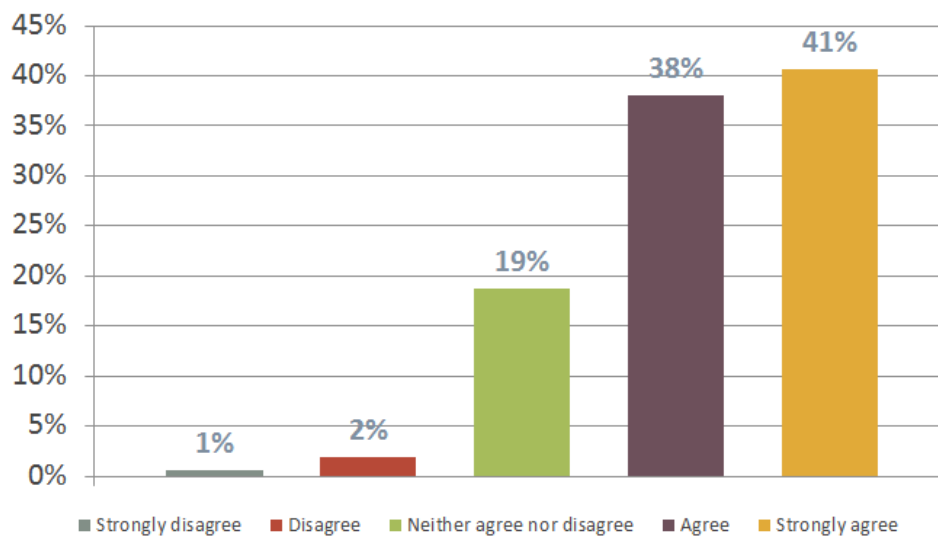
Q6 in Solar Survey,  
Q6 in EE Survey

**Figure 3. How knowing the value of a storage system might affect sales**

### *Providing Home Energy Needs*

Almost 4 of every 5 respondents, 79 percent, said they were motivated to get a battery backup system for the home, knowing that a battery backup system combined with a solar system can provide a home's basic energy needs.

Only 3 percent said they would not be inclined to purchase one while 19 percent said they were open to the possibility.



N= 155

Q7 in Solar Survey,  
Not in EE Survey

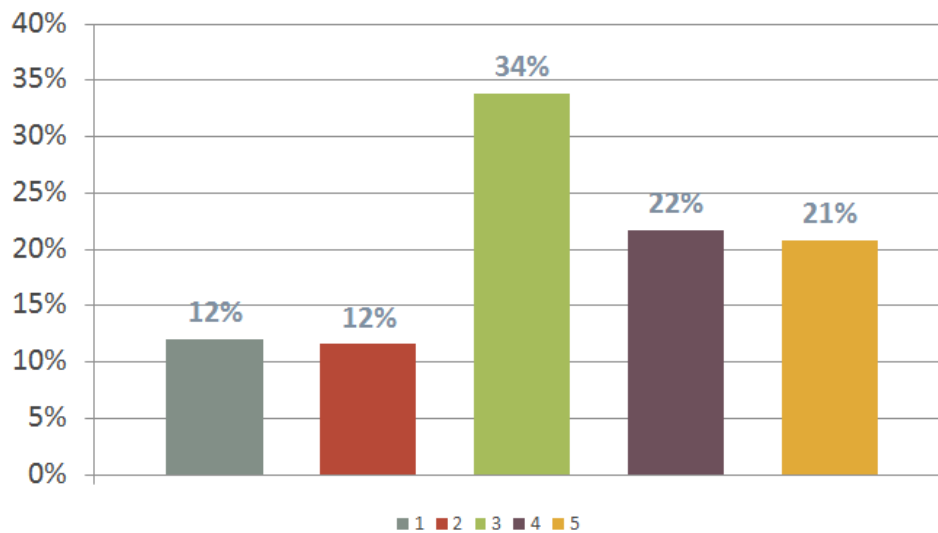
**Figure 4. How basic energy needs might spur sales**



### Natural Disasters

More than 4 of 10 respondents, 43 percent, said they had concerns about safety due to power outages arising from earthquakes or other natural disasters, with 21 percent saying they “strongly” agree with the suggestion that they harbor such worries.

About a quarter of respondents, 24 percent, said they had no such concerns, with the remainder, 34 percent, expressing no feelings either way.



N= 207

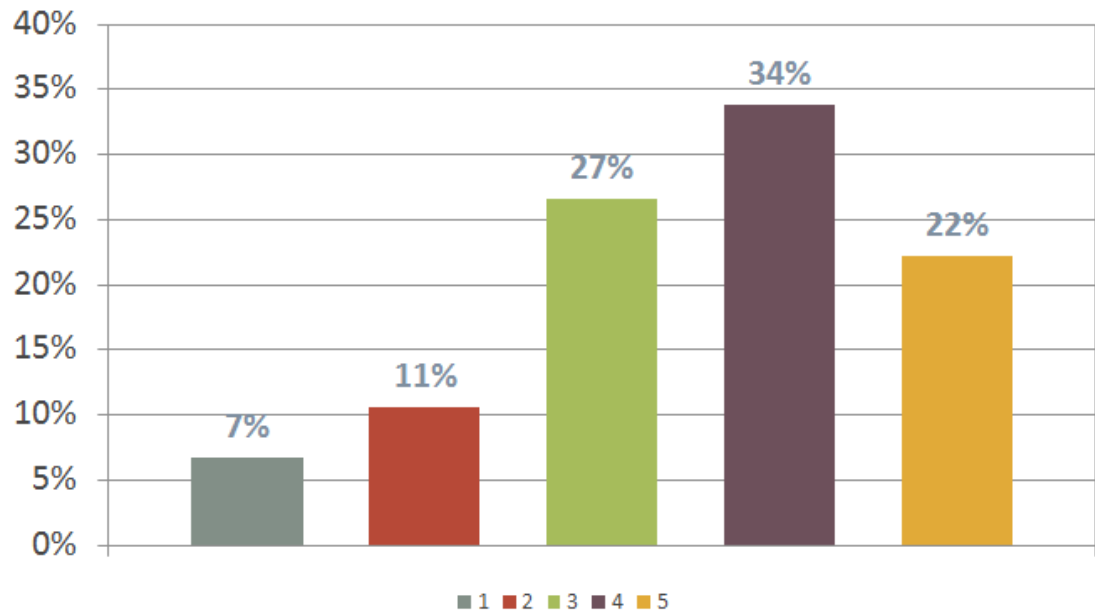
Q8 in Solar Survey,  
Q8 in EE Survey

**Figure 5. Worrying about natural disasters and energy disruption**

### *Spending a Little More*

More than half of respondents, 56 percent, agreed that they would be willing to spend “a little more” on products and services that decrease environmental impact, with 22 percent of respondents feeling “strongly” that they would.

Only 7 percent felt “strongly” that they wouldn’t.



N= 207

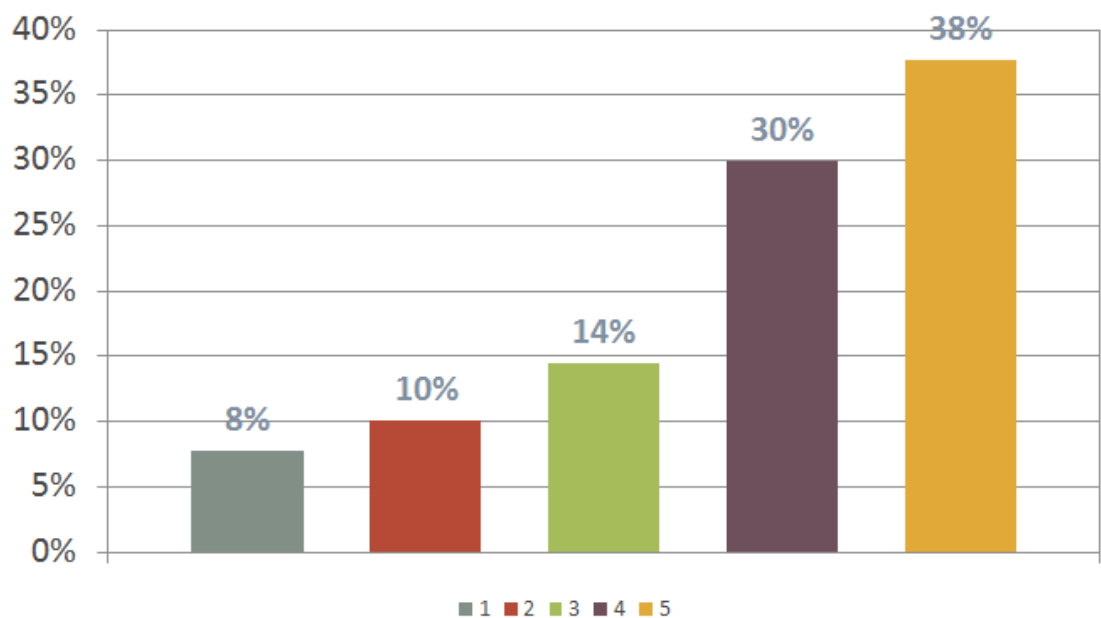
Q9 in Solar Survey,  
Q9 in EE Survey

**Figure 6. Willingness to spend more to help the environment**

### Personal Control

More than two thirds of all respondents, 68 percent, said that having more personal control over their energy production and usage was important to them, and 38 percent of all respondents indicated that they felt “strongly” about it.

Fewer than 1 in 5 of all respondents, 18 percent, disagreed with the idea that more personal control was important to them.



N= 207

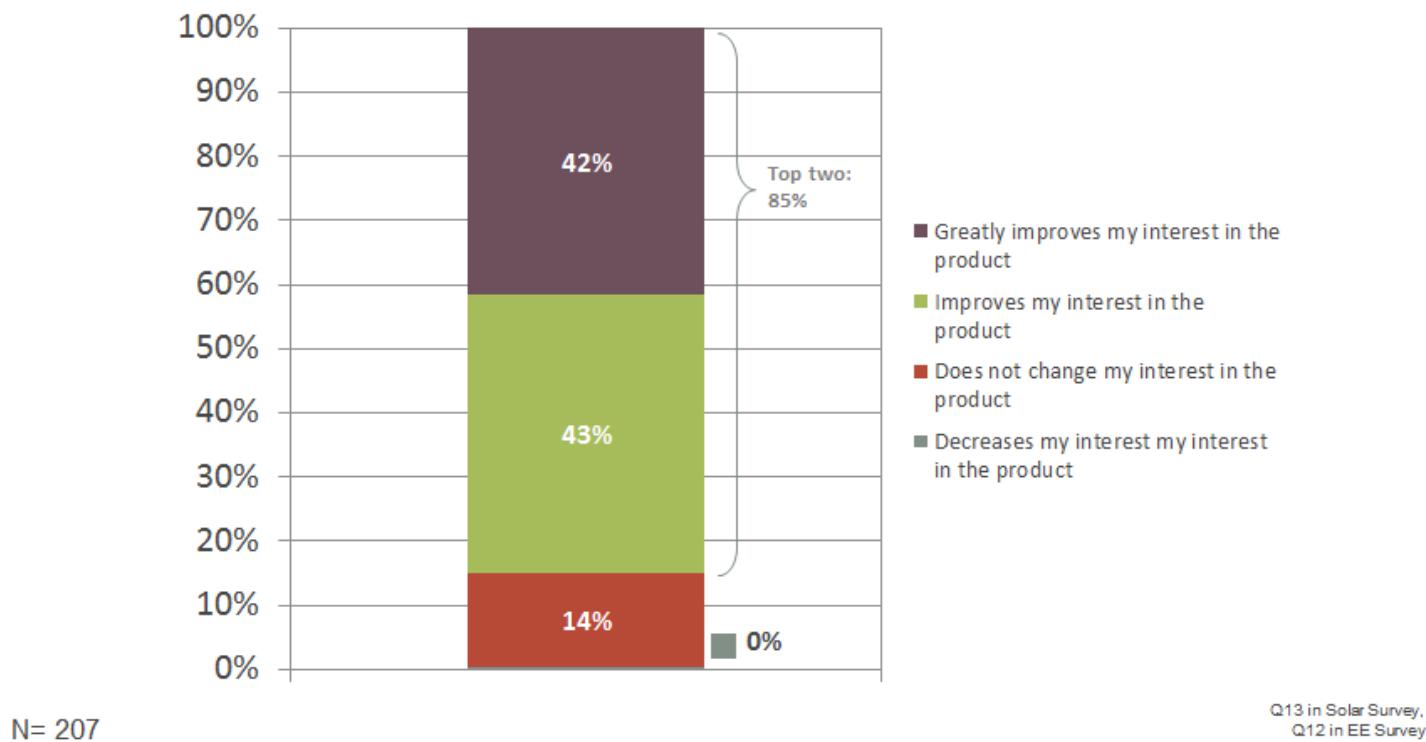
Q10 in Solar Survey,  
Q10 in EE Survey

**Figure 7. Importance of controlling personal energy production**

### Limited-Time Rebates

A vast majority of respondents, 85 percent, said their interest in a storage battery improved, knowing that they could lower their energy costs through limited-time rebates from some states and utility districts.

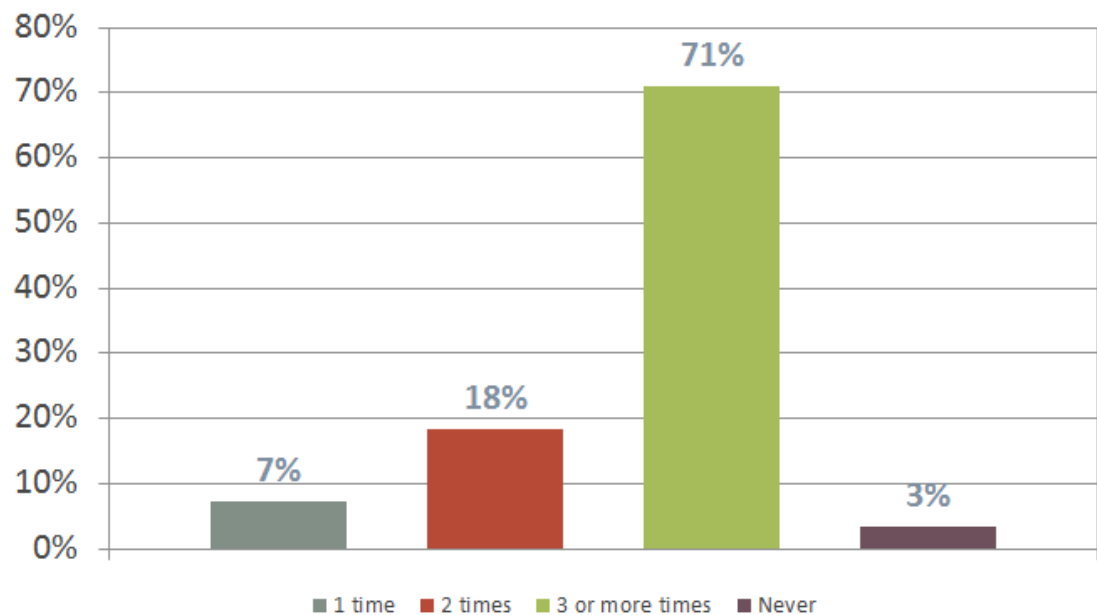
Another 14 percent said the rebates would not change their interest in battery storage.



**Figure 8. The possible impact of rebates on purchases**

### Experiencing Energy Outages

Almost all survey respondents, 96 percent of them, said they had experienced at least one energy outage in the last five years, with 71 percent saying it had happened three times or more. The results showed that 18 percent had experienced outages twice in five years, and 7 percent said it happened once.



N= 207

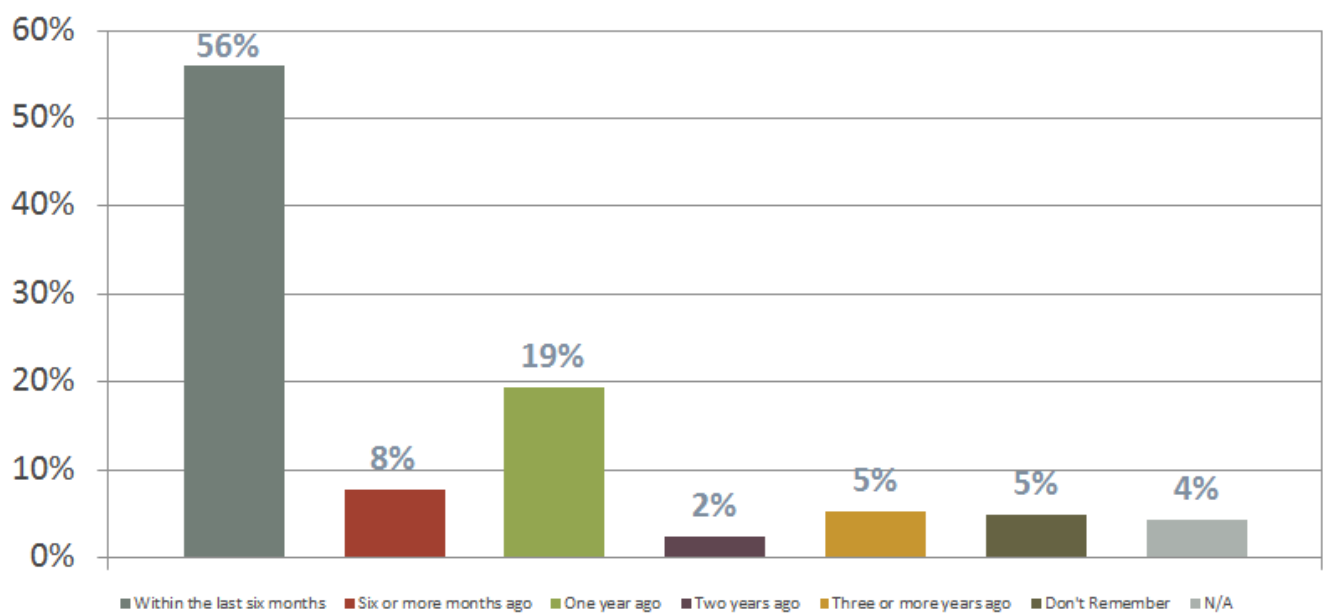
Q15 in Solar Survey,  
Q14 in EE Survey

**Figure 9. Energy disruptions within the last five years**

### *Timeliness of Outages*

More than half of respondents, 56 percent, said they last lost power within the previous six months. Another 27 percent said the power last went out within the past year.

The remaining respondents said the power went out two years ago or more or they didn't remember when it happened.



N= 207

Q16 in Solar Survey,  
Q15 in EE Survey

**Figure 10. Frequency of recent energy outages**

### *Customer Worries*

When power goes out the problem cited by the largest number of people, 38 percent, was the condition of food in the refrigerator. The other leading areas of concern were heat (18 percent), comfort/security (14 percent), water (8 percent) and light (8 percent).

Other concerns raised were the loss of communications (5 percent), the sump pump (1 percent) and air conditioning (1 percent). Three percent of people said they were concerned about “everything.”

Concern	Count	%
Food in the fridge	78	38%
Heat	38	18%
Comfort and security	30	14%
Water	17	8%
Light	16	8%
Communication	11	5%
Everything	6	3%
None	4	2%
Sump pump	3	1%
AC	2	1%
Not being there to turn on/reset	1	0%
We have a generator	1	0%

N= 207

Q17 in Solar Survey,  
Q16 in EE Survey

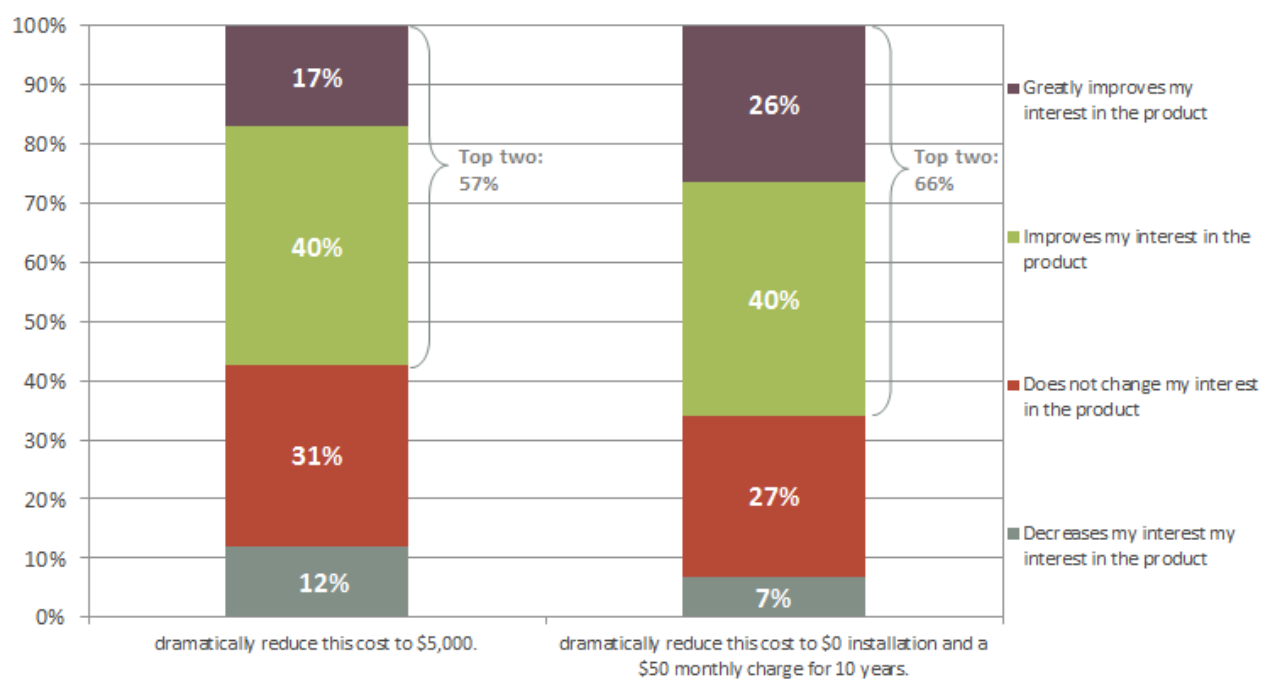
**Figure 11. Issues of concern caused by an outage**



### Impact of Rebates

Customer interest in purchasing a battery storage system that costs \$20,000 to manufacture and install rises with the possibility that government rebates can reduce the price.

A rebate program that would reduce the cost to \$5,000, intrigued more than half of respondents, 57 percent, with 40 percent saying their interest improves and 17 percent saying their interest “greatly” improves. A program with no installation costs and a \$50 monthly fee for 10 years improved the interest in such a program of 40 percent of respondents and greatly improved the interest of 26 percent.



N= 207

Q18 in Solar Survey,  
Q17 in EE Survey

**Figure 12. Rebate programs and customer interest**

## Conclusion

With recent extreme weather events fresh in their minds, those customers that responded to the survey appear overwhelmingly aware and interested in the benefits that energy storage combined with PV can provide including back-up power, environmental benefits, and financial benefits. Well-placed incentives and financing options appear to drive purchase interest for early adopters of battery systems. To deliver on this potential, the industry must continue to help package new storage technologies in ways that best balance the needs of the customer with the broader societal benefits that storage incentive programs seek to drive.

## 4.2 PV Variability Analysis

### Summary

This is one of three final report sections completed by a UC Berkeley team of students and faculty in the Energy and Resources Group. It covers efforts from Task 4.2. Our specific focus is on understanding distribution system impacts and the opportunity for creating value by incorporating storage into the CAISO's dispatch process.

This particular report documents our efforts to study the impact of distributed PV portfolios on grid operations and planning.

In Part 1 (Section 4.2.1), we develop and test a statistical model to predict the worst case regulation and load following requirements resulting from large numbers of distributed photovoltaic (PV) systems. Our model accounts for geographic autocorrelation between generation from closely located systems, and regulation and load following requirements from each system during one time period are modeled as normal distributions whose parameters depend on a "volatility state" that is estimated within the model. The estimated volatility state is loosely interpretable as representing the size distribution and overhead speed of clouds. The model is fit using 1-minute resolution generation data from 100 PV systems in the California Central Valley or the Los Angeles Coast. We also make the volatility state of each system dependent on widely available 15-min generation data from nearby PV systems.

The model developed in Part 1 serves as the driver for a range of simulation studies developed in Part 2 (Section 4.2.2). In particular we predict the load following and AGC requirements to follow variable generation from solar photovoltaics (PV) in California for 1) different arrangements of PV in California, e.g., centralized versus distributed 2) different sizes of balancing areas, e.g., balancing of each system versus balancing the aggregate signal the whole state, and 3) different total amounts of Solar PV in California. This model is uniquely able to account for geographic diversity and stochastic volatility of generation from any set PV panels within an region. We find that if PV is distributed on rooftops throughout the state and balancing is done across the entire aggregation of systems, the largest regulation and load following requirement totals roughly 900 MW for 12 GW of PV. On the other hand, PV is centralized (i.e. utility scale), the requirement doubles to 1800 MW. We also find that there is a very large penalty to local balancing (rather than balancing at the state level), with a maximum total requirement of roughly 6500 MW of balancing capacity for 12 GW of PV.

#### 4.2.1 Modeling the Regulation and Load Following Requirements of Distributed Photovoltaics

**4.2.1.1 Introduction** Grid-connected solar photovoltaic (PV) systems are rapidly increasing in number. As of October 2013 Over 240,000 PV systems are installed in the US, amounting to over 5.5 GW of total rated capacity. Most of these installations are concentrated in cities in the West and the Desert Southwest. For example, over 2.8GW (170,000 installations) of these are in California and Arizona [1]; 550MW in highly urbanized counties of Los Angeles, Riverside, San Diego, and Santa Clara. California's goals for the next decade are even more ambitious: Combined initiatives include subsidies for 3 GW of distributed PV by 2016 [2], and future goals aim to increase this amount four-fold to 12 GW of distributed generation (mostly PV) by 2022 [3].

This rise in distributed PV coincides with concerns that intermittent generation from wind and solar will make it increasingly difficult to balance supply and demand and lead to increased costs for maintaining reliability. Specifically, additional, uncontrolled variability from wind and solar may require additional, fast-responding generators that can take many years to go through permitting and construction. Recent grid planning studies have attempted to quantify the requirements of future variability [4, 5, 6, 7]; however, assessing the amount of fast-responding resources that distributed PV will require from power systems remains difficult. Power systems planners are concerned with high-impact, low-probability events, and there is a limited amount of high resolution data from which to sample for simulations. Robust statistical models of variability from renewables are needed to accurately assess the future needs of the power system without relying on expensive, conservative assumptions.

Many studies have shown geographic autocorrelation in PV variability, meaning that closely located PV systems are more likely to fluctuate in the same direction than those far apart [8, 9, 10, 11, 12]. Most studies directly relate PV generation signals or clear sky indices from neighboring sensors using correlation coefficients. All of these studies show that the correlation of variability between two systems decreases as the distance between them increases. See Mills and Wiser [11] or Lave et al. [8] for a complete literature review of this field.

Murata [10] shows that the predicted correlation coefficient between all pairs of system in a network can be used to find the variance of the aggregate generation from the entire network. However, actual estimation of the distribution of likely generation at a given time remains difficult due to the changing shape of the distribution as signals are added, as shown in Murata's and Mills and Wiser's studies. Specifically, the distribution of variable signals from PV exhibits a high kurtosis at individual sites (a distribution with a high peak and fat tails); the kurtosis decreases when generation from multiple sites are added together [11]. Previous work on geographic auto-correlation of PV generation predicts the mean and variance of variability, which are sufficient parameters to characterize only a few distribution shapes (most notably, the Gaussian normal distribution). Given that the shape of fluctuations (even after aggregation) is unidentified and specifically non-Gaussian, mean and variance are insufficient.

The inability to predict the distribution diminishes the usefulness of statistical models for risk assessment of solar PV in power systems. Power system planners and operators are concerned with high-impact, low probability events, thus the tail of a distribution is arguably the most important aspect. However the tail is greatly affected by the unpredictable shape. Recent planning studies have not used statistical models for solar PV generation, they instead repeatedly substitute existing short time-scale data into simulations [7].

One aspect of PV generation that leads to changing distribution shape is the existence of different cloud

regimes that affect variability, i.e., PV generation signals will be highly volatile in partly cloudy times with fast wind speeds (resulting in the fat distribution tail) and much less volatile during sunny times or fully overcast times (resulting in the high peak). Recent studies have attempted to identify different cloud regimes and fit separate models for each of them. Lave and Kleissl condition their geographic auto-correlation parameters on cloud size and speed predictions from numerical atmospheric models [13, 14]. Hoff and Perez condition their variability and geographic auto-correlation parameters on the spatial variability of satellite predicted horizontal irradiance [15, 16]. Hummon et al. use variability “classes” to simulate the effects of different cloud regimes on PV generation time series for simulation purposes [17]. Reno and Stein put forward the use of a “Variability Index” that classifies days using the standard deviation of the cloud cover index [18].

In this section, we present, fit, and validate a model for predicting the largest likely requirements for regulation and load following to follow a variable generation signal from distributed PV systems. Our model conditions volatility on a hidden “volatility state” that defines the variance of the PV generation signal as well as the geographic autocorrelation function among nearby systems. Our model differs from previous work by endogenously estimating volatility states as latent variables with distributions dependent on external data. By using this method we can parameterize distributions of variable signals from solar PV as gaussian mixtures that are consistent with empirical data even in the tails. Estimation is completed using expectation-maximization (EM) of a hidden Markov model (HMM) as described in our previous work [19]. Our model is unique in that it

1. **predicts operational requirements** for following a variable signal, as defined in Makarov, 2009 [6].
2. **uses generation data** for fitting and validation, measured at over 100, closely located, residential PV systems in California.
3. **predicts a distribution shape** by conditioning on endogenously estimated latent variables that allow the predicted shape to be a Gaussian mixture.
4. **is conditioned on widely available external data** and thus can be applied in external geographic areas while accounting for their unique sky conditions.

**4.2.1.2 Data and Processing** Instantaneous voltage and current measurements were taken each minute from the inverters of over 150 residential and commercial PV installations owned and maintained by SolarCity® and provided to us under the terms of a nondisclosure agreement. All of the installations are contained within 256km<sup>2</sup> areas of either Los Angeles (LA), or the California Central Valley (CV), monitored from mid-June to the end of August 2012. We selected systems from among many using an algorithm that combines (1) quota sampling for distances between pairs of locations and (2) geographically random sampling of pairs of sites.

Data were collected at the inverters for each system and sent to servers via ZigBee devices. Communication limitations resulted in large gaps in data collection. We use a subset of total data that contains 30 days of consistently recorded generation from 39 inverters in LA and 55 in CV. Figure 3 shows rough locations for systems in this final dataset, referred to as the “fleet”.

Instantaneous voltage and current at 15 minute resolution are available under the same non-disclosure agreement for over 6000 systems from January 2011 to late September 2012.

**4.2.1.2.1 Decomposition of variable signal into fast-response requirements** We use the method put forward in Makarov 2009 [6] to estimate two impacts of a variable generation signal on power system operations: 1) a load following signal, which represents the generation procured at each clearing of a real time market (5-minute intervals) to follow a variable net load signal; and 2) an Automatic Generation Control (AGC) signal, which represents the generation needed every minute to balance net-load and variable generation beyond the real time market.

Equations 1 and 2, reproduced from Makarov 2009, are used to estimate the total dispatched generation the hourly markets,  $D_{ha}(t)$ , and total dispatched generation following the load following markets,  $D_{5min}(t)$ ; where  $L(t)$  is a variable load signal,  $avg_T(\cdot)$  is an operator that produces averages over discrete intervals of duration  $T$ ,  $\mathfrak{R}_T(\cdot)$  is an operator which adds ramps of duration  $T$  between interval averages, and  $\epsilon$  represents forecast error. Examples of  $D_{ha}$  and  $D_{5min}$  are shown in panel A of Figure 1 as calculated for the net load of one solar installation.

$$D_{ha}(t, L) = \mathfrak{R}_{20min}\{\text{avg}_{60min}(L(t)) + \epsilon_{ha}\} \quad (1)$$

$$D_{5min}(t, L) = \mathfrak{R}_{5min}\{\text{avg}_{5min}(L(t)) + \epsilon_{5min}\} \quad (2)$$

Generation purchased in the load following market is required to meet the demand stipulated in  $D_{5min}$  while other generation is already meeting demand stipulated in  $D_{ha}$ . Thus, load following requirements,  $G_{lf}$ , are the difference between  $D_{ha}$  and  $D_{5min}$ , shown in Equation 3; where  $j$  signifies positive integer minutes, and  $\tau$  is a constant of 5 minutes representing the interval for the load following market. Generation used in AGC is required to make up the difference between the total net load and the demand stipulated in  $D_{5min}$ , shown in Equation 4.  $G_{lf}$  and  $G_{agc}$  are shown in panels B and C of Figure 1, both requirements can be negative, signifying that generators are required to quickly lower power generation to prevent an oversupply of electricity.

$$G_{lf}(\tau t, L) = D_{5min}(\tau t, L) - D_{ha}(\tau t) \quad (3)$$

$$G_{agc}(t, L) = L(t) - D_{5min}(t, L) \quad (4)$$

Both decompositions result entirely from linear operations, which provides a helpful property for our analysis shown in Equation 5: the requirement to follow the sum of multiple load signals at a given time is the same as the sum of the requirements to follow each load signal at that time.

$$G_{agc,lf}\left(t, \sum_{i=1}^n L_i\right) = \sum_{i=1}^n G_{agc,lf}(t, L_i) \quad (5)$$

This work focuses on estimating the impacts of variability only by setting  $\epsilon$  equal to one. We focus on variability alone because we seek to characterize the basic impacts of the underlying meteorological process without assumptions specific to a particular forecast model; as we discuss in the conclusions forecast error models can easily be added to our approach. However this means that for the present study our estimates are low.

**4.2.1.2.2 Estimation of Clear Sky Signal** Our statistical model relies on the knowledge of a “clear sky signal” which represents predictable generation that would have occurred in absence of cloud cover.

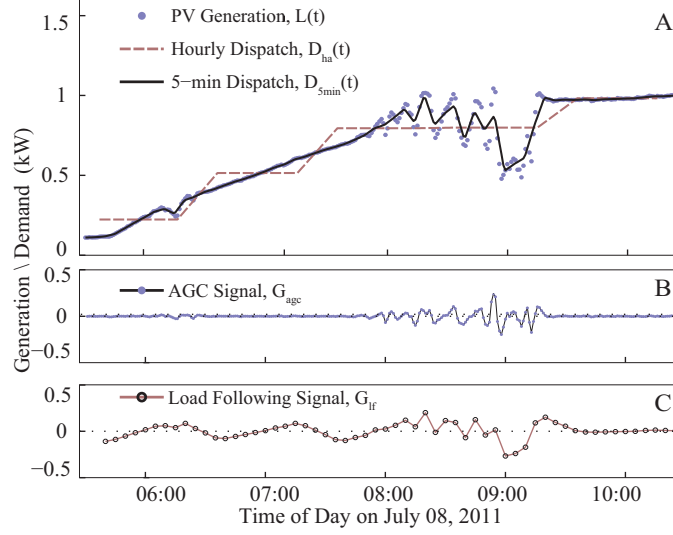


Figure 1: Decomposition example of a variable net-load signal (PV generation) into operational requirement signals. (A) Net load along with total generation after market clearing of the hour-ahead and real-time markets, (B) the resulting AGC requirement signal, (C) the resulting load following signal.

Solar-earth geometry drives the clear-sky signal, and this geometry is predictable given the time of year, and the location, tilt, azimuth, and derated capacity of a PV system. However, our metadata on the geometry and total rated capacity of panels connected to each inverter are not always accurate. In addition humidity, outdoor temperature, and shading also affect generation in ways not predicted by geometry.

To overcome these disparities, we create an empirically corrected clear sky signal. First, we find the deviation of actual measurements from the clear sky predicted measurements using the solar-earth geometry model described in [20]. For each deviation reading, we find the 95<sup>th</sup> percentile of all deviations during the same daily interval over the course of a four-week window. The 95<sup>th</sup> percentile is used in attempt to exclude any low observations due to cloud-cover (common) as well and any high observations due to cloud reflection (much less common). The resulting signal is then smoothed using a 2 hour moving average and added to the predictions from the solar-earth geometry model. This method both corrects for errors in the system metadata as well as allows for periodic shading from buildings, provided the effect is long enough not to be smoothed out. It also occasionally accounts for periodic morning/evening cloud cover that persists for more than four weeks. Figure 2 shows 1-minute resolution generation for one day along with the clear sky signal before and after smoothing.

**4.2.1.2.3 Volatility heuristics from 15-min measurements** We condition our model on 15-minute generation measurements taken at systems nearby to those in our fleet, referred to as “knots,” shown in Figure 3. To choose knots, we subset our study regions into 2km grids; for each grid cell, the system with complete data that is closest to the centroid of each grid cell and is not part of the 1-min dataset is chosen as a knot.

The inputs to our model are volatility heuristics from 15-min measurements, which are calculated using a moving standard deviation from a moving average. First, 15-min measurements are normalized using the clear sky signal, so that regardless of the rated capacity of a system or the time of day the signal should be



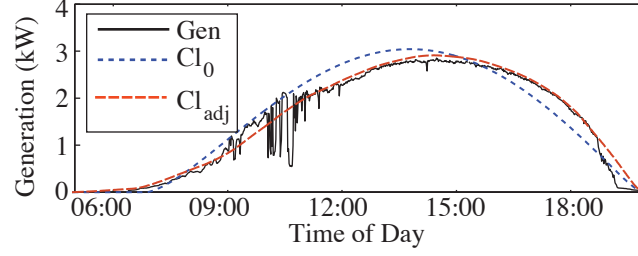


Figure 2: 1-minute resolution generation and two clear sky signals calculated for one day in Los Angeles.  $Cl_0$  is found using solar-earth geometry and system geometry from metadata,  $Cl_{adj}$  is the adjusted clear sky signal following smoothing.

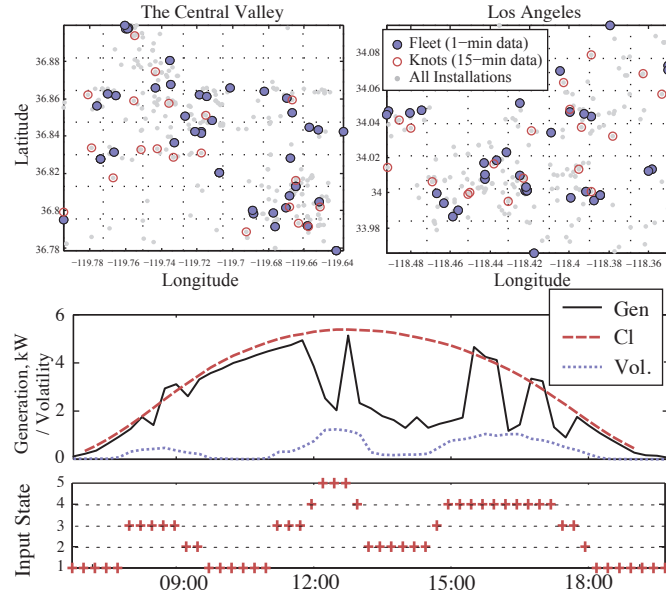


Figure 3: Panels A and B show the locations of knots and the fleets of systems for which we fit and validate the model. Panel C shows 15-minute resolution generation, clear sky generation, and the moving variance from one knot. Panel D shows the input state of this knot based on the moving variance in panel C.

remain between 0 and about 1.3. A moving average is taken of the normalized signal using a centered window shown in Equation 6. A moving standard deviation of the signal from the moving average is then calculated as shown in Equation 7. The moving standard deviation is placed into 5 bins with lower bin edges defined using exponential intervals to be  $0, e^{-3.5}, e^{-2.83}, e^{-2.16}$ , and  $e^{-1.5}$ . The resulting time series of bin placements is used as the heuristic. Panel C of Figure 3 shows one day of generation from a knot along with the moving variance of its generation; panel D shows and the resulting volatility heuristics for this day.

$$\mu(t) = \frac{1}{m+1} \sum_{i=(t-\frac{m}{2})}^{t+\frac{m}{2}} \frac{G(t) - Cl(t)}{\max_{t \in DOY}(Cl(t))} \quad (6)$$

$$\sigma_v(t) = \frac{1}{m+1} \left[ \sum_{i=(t-\frac{m}{2})}^{t+\frac{m}{2}} \left( \frac{G(i)}{cl(i)} - \mu(i) \right)^2 \right]^{\frac{1}{2}} \quad (7)$$

**4.2.1.3 Statistical Model** The premise of our statistical model is that PV generation is more effectively modeled by accounting for periods of time when generation is highly volatile (e.g., partly cloudy) or non-volatile (e.g., sunny). We do not observe these patterns directly, and instead estimate them endogenously as latent states, referred to as volatility states. By identifying volatility states within which the distribution of requirements is Gaussian, our model is able to predict not only the variance but also the distribution of the resulting requirements.

A graphical representation of our model is shown in Figure 4 and described in this section. Figure 4 depicts a directed acyclic graph, where shaded nodes represent observed variables and unshaded nodes represent unobserved variables; arrows indicate dependence, where the distribution of the “to” node is dependent on information in the “from” node. The AGC or load following requirement for each system at time  $t$  is denoted  $Y(t) \in \mathbb{R}^n$ , where  $n$  is the number of systems.  $Y(t)$  is modeled as a multivariate Gaussian, shown in Equation 8. The mean of  $Y(t)$  is taken to be the requirement needed to follow the clear-sky signal at time  $t$ ,  $\vec{Y}_{CL}(t) \in \mathbb{R}^n$ .  $\vec{v}$  is a mean zero random variable that represents variability other than that predicted by the clear sky signal, defined in Equation 9. It is scaled by the hourly maximum of the clear sky signal to account for the fact that clouds cause greater variability when the generation from the sun is high.

$$\vec{Y}(t) = \vec{Y}_{CL}(t) + \vec{v}(t) \cdot \max_{t \in HE} CL(t) \quad (8)$$

$$\vec{v}(t) \sim MVG(0, \Sigma_Y(\vec{v}_t; \phi)) \quad (9)$$

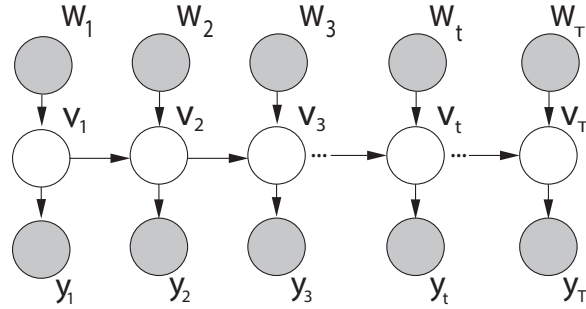


Figure 4: Statistical model represented as a directed acyclic graph

The covariance matrix of  $\vec{v}(t)$  is denoted  $\Sigma_Y$  and varies with time, defined in Equation 10.  $\Sigma_Y$  is conditionally dependent on the volatility state of each system,  $v_t = \{1, 2, \dots, M\}^n$ , and on an exponential geographic autocorrelation function defined in Equation 11.  $\vec{v}_t$  is an  $n$ -valued multinomial random variable with  $M$  possible values, where the  $i^{th}$  element represents the volatility state of system  $i$  at time  $t$ .  $\sigma^2 = \{\sigma_1^2, \sigma_2^2, \dots, \sigma_M^2\}$  is a set of variances whose  $i^{th}$  element corresponds to the variance in the  $i^{th}$  volatility state.

$$\Sigma_{ij}(\vec{v}_t; \vec{\sigma}, \phi) = \begin{cases} \sigma_{v_{t,i}}^2 & i = j \\ \sigma_{v_{t,i}} \sigma_{v_{t,j}} \rho(v_{t,i}, v_{t,j}; d_{i,j}, \phi) & i \neq j \end{cases} \quad (10)$$

An exponential geographic autocorrelation function is used, as defined in Equation 11, where  $\phi = \{\mathbf{a}, \tau\}$ .  $a_{m,n} \in [-1, 1]$  represents the correlation between two separate but identically located system where one is in state  $m$  and the other is in state  $n$ .

$$\rho(m, n; d_{i,j}, \phi) = a_{m,n} \cdot \exp \{-d_{i,j}/\tau_{m,n}\} \quad (11)$$

$a_{m,n}$  may be less than one signifying that there is variability between two systems that are located at the same installation site but metered separately. Though they are adjacent, two systems at the same site occupy different locations. Imperfect correlation for these systems may result from irregular shading of module strings across the spatial coverage of the site.  $\tau_{m,n}$  is a range parameter representing the distance over which correlation decays by 63%. A separate exponential function is fit for each possible pairing of volatility states, denoted  $m$  and  $n$ , thus we fit  $\frac{(M+1)(M)}{2}$  different autocorrelation functions for a model with a total of  $M$  volatility states.

The only latent random variables in our model are the volatility states. These states are conditional on the volatility state of the previous time-step at the same system,  $v_{t-1,n}$ , and on the input heuristic from the closest knot at time  $t$ ,  $W_{t,m}$  (calculated from 15-min resolution generation data as described in Section 4.2.1.2.3). Equation 12 shows the distribution of  $\vec{v}_t | \{\vec{W}_t, \vec{v}_{t-1}\}$  as a set of transition matrices for Markov chains where there is one transition matrix,  $\mathbf{A}^{(k)} \in \mathbb{R}^{M \times M}$ , for each input heuristic  $k \in \{1, 2, \dots, K\}$ . Equation 13 describes each transition matrix.

$$\mathbb{A} = \{\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \dots, \mathbf{A}^{(K)}\} \quad (12)$$

$$\mathbf{A}_{i,j}^{(k)} = P(v_{t,n} = j | v_{t-1,n} = i, W_{t,m} = k) \quad (13)$$

In total, the model contains  $M + 2\frac{(M)(M+1)}{2} + M(M-1)K$  parameters

1.  $M$  variances, one for each volatility state
2.  $2\frac{(M)(M+1)}{2}$  parameters in the auto covariance function, an  $a$  and a  $\tau$  parameter for each pair of volatility states.
3.  $M(M-1)K$  parameters, one  $M \times M$  transition matrix for each of  $K$  input states.

**4.2.1.3.1 Estimating the model** We estimate two separate models, one for load following requirements and one for AGC requirements. We test each model's performance between 3-8 volatility states. Cross validation is applied over time by randomly selecting 75% of the days and fitting the model with these — we refer to these data as the “model” data. The remaining 25% of the data, referred to as the “test” data, are used for validation.

Parameters are estimated in two stages: the first stage estimates the entire model assuming no geographic autocorrelation, the second stage estimates autocorrelation parameters for each pair of volatility states given the output from stage 1.

**Stage 1:**  $\vec{v}_t$ ,  $\vec{\sigma}$ , and  $\mathbb{A}$  are estimated using Expectation-maximization (EM) of a hidden Markov model. EM iteratively executes the following steps: first, it chooses the maximum likely parameters of the model given the expected value of the volatility states at each time, then it recalculates the expected value of the

volatility states given the updated model parameters, it then repeats until convergence. Because EM is a gradient ascent method, is not guaranteed to find the global maximum but it will find a local maximum.

EM provides us with expected values of volatility states, defined in Equation 14, where  $\gamma_{i,m,t}$  is the expected value of an indicator for whether system  $i$ , is in volatility state  $m$ , at time  $t$ .

$$\gamma_{i,m,t} = E[\mathbb{I}\{v_{i,t} = m\}] \quad (14)$$

**Stage 2:** Equation 15 describes a weighted correlation coefficient  $\rho_{i,j,m,n}$  between the AGC or load following signal for system  $i$  when in state  $m$  and for system  $j$  when in state  $n$ . The covariances and variances used to find each correlation coefficient is weighted by the probability each system is in their respective volatility state,  $p(v_{t,i} = m, v_{t,j} = n) = \gamma_{i,m,t}\gamma_{j,n,t}$

$$\rho_{i,j,m,n} = \frac{\frac{\sum_{t=1}^T y_{t,i} y_{t,j} \gamma_{i,m,t} \gamma_{j,n,t}}{\sum_{t=1}^T \gamma_{i,m,t} \gamma_{j,n,t}}}{\sqrt{\frac{\sum_{t=1}^T \gamma_{i,m,t} \gamma_{j,n,t} y_{t,i}^2}{\sum_{t=1}^T \gamma_{i,m,t} \gamma_{j,n,t}}} \sqrt{\frac{\sum_{t=1}^T \gamma_{j,n,t} \gamma_{i,m,t} y_{t,j}^2}{\sum_{t=1}^T \gamma_{j,n,t} \gamma_{i,m,t}}}} \quad (15)$$

The auto-correlation parameters,  $a$  and  $\tau$ , are fit using weighted least squares from the correlation for each pair of sites. Each correlation coefficient,  $\rho_{i,j,m,n}$ , is weighted by the expected number of time-steps that the systems  $i$  and  $j$  are in states  $m$  and  $n$  respectively, shown in Equation 16. Equation 17 shows the weighted objective function for fitting the autocorrelation parameters.

$$w_{i,j,m,n} = \sum_{t=1}^T \gamma_{i,m,t} \gamma_{j,n,t} \quad (16)$$

$$\{a_{mn}, \tau_{mn}\} = \min_{a, \tau} \sum_{i,j=1}^N \left( a \cdot e^{\frac{-d_{ij}}{\tau}} - \rho_{i,j,m,n} \right)^2 w_{i,j,m,n} \quad (17)$$

**4.2.1.4 Paramater Estimation Results** Tables 1 and 2 display the log-likelihood of the model data and the test data for the AGC model and the load following model. For both the model and the test data, the log-likelihood increases monotonically with the number of possible volatility states. This implies that additional states increase the predictive power of the model without overfitting noise in the model data. However, there is a sharp decreases in the rate of increase of the log-likelihood with the addition of states beyond 5, signifying that additional states are not adding much predictive power.

Tables 1 and 2 also compare variance parameters fit for each volatility state in each model. Here the volatility states are ordered so that state 1 has the lowest variance and state  $M$  has the greatest. The  $f$  column shows the fraction of system-time spent in each state. For  $M > 5$ , some states are rarely encountered, this is a sign that model has too many states and that the rarely encountered ones are adding very little usefulness. The inclusion of rarely encountered states begins at  $M = 6$ , which is also where the log-likelihood stops increasing. For the remainder of our analysis, we use models with 5 states for both the AGC and load following requirements.

Table 1: Hidden Markov model estimation results for the load following requirement signal. The log-likelihoods for models fit with 3 to 8 states are reported using both the model data and the test data. The variance estimated for each hidden state in each model,  $\sigma$ , is also reported along with the total fraction of system time spent in each state,  $f$

	<b>M = 3</b>		<b>M = 4</b>		<b>M = 5</b>		<b>M = 6</b>		<b>M = 7</b>		<b>M = 8</b>	
ELL Model	4.581e+06		4.705e+06		4.735e+06		4.737e+06		4.745e+06		4.745e+06	
ELL Test	1.488e+06		1.527e+06		1.534e+06		1.535e+06		1.537e+06		1.537e+06	
<b>State #</b>	$\sigma^2$	$f$	$\sigma^2$	$f$	$\sigma^2$	$f$	$\sigma^2$	$f$	$\sigma^2$	$f$	$\sigma^2$	$f$
1	1e-05	0.64	3e-06	0.33	3e-06	0.31	3e-06	0.31	3e-06	0.3	3e-06	0.3
2	0.0003	0.22	2e-05	0.38	2e-05	0.37	2e-05	0.36	2e-05	0.33	2e-05	0.33
3	0.02	0.13	0.0006	0.18	0.0002	0.15	0.0002	0.14	8e-05	0.13	8e-05	0.13
4	-	-	0.02	0.11	0.003	0.11	0.002	0.11	0.0006	0.12	0.0006	0.12
5	-	-	-	-	0.03	0.065	0.02	0.0065	0.005	0.076	0.005	0.076
6	-	-	-	-	-	-	0.03	0.063	0.02	0.004	0.02	0.00082
7	-	-	-	-	-	-	-	-	0.04	0.047	0.02	0.0032
8	-	-	-	-	-	-	-	-	-	-	0.04	0.047

Table 2: Hidden Markov model estimation results for the load following requirement signal. The log-likelihoods for models fit with 3 to 8 states are reported using both the model data and the test data. The variance estimated for each hidden state in each model,  $\sigma$ , is also reported along with the total fraction of system time spent in each state,  $f$

	<b>M = 3</b>		<b>M = 4</b>		<b>M = 5</b>		<b>M = 6</b>		<b>M = 7</b>		<b>M = 8</b>	
ELL Model	7.429e+05		7.525e+05		7.566e+05		7.57e+05		7.584e+05		7.584e+05	
ELL Test	2.073e+05		2.106e+05		2.12e+05		2.12e+05		2.124e+05		2.124e+05	
<b>State #</b>	$\sigma^2$	$f$	$\sigma^2$	$f$	$\sigma^2$	$f$	$\sigma^2$	$f$	$\sigma^2$	$f$	$\sigma^2$	$f$
1	2e-05	0.52	2e-05	0.45	1e-05	0.37	1e-05	0.37	6e-06	0.11	6e-06	0.11
2	0.0005	0.26	0.0002	0.22	7e-05	0.22	7e-05	0.22	2e-05	0.31	2e-05	0.31
3	0.01	0.22	0.002	0.21	0.0007	0.17	0.0006	0.16	9e-05	0.18	9e-05	0.18
4	-	-	0.02	0.13	0.005	0.17	0.003	0.0027	0.0007	0.15	0.0007	0.15
5	-	-	-	-	0.03	0.069	0.005	0.17	0.003	0.0034	0.003	0.003
6	-	-	-	-	-	-	0.03	0.074	0.005	0.17	0.005	0.17
7	-	-	-	-	-	-	-	-	0.03	0.073	0.01	0.0035
8	-	-	-	-	-	-	-	-	-	-	0.03	0.072

Tables 3 and 4 show estimates of the autocorrelation parameters,  $a_{m,n}$  and  $\tau_{m,n}$ , for AGC and load following respectively. For both requirements,  $a$  (the correlation at a distance of zero) generally increases as volatility state increases. This result is intuitive, clouds are geographically autocorrelated and create large magnitude variability, thus we expect that as the variance increases, so will the amount of variability attributed to geographic auto-correlation.  $\tau$  exhibits a non-monotonic relationship with volatility state, where the decay range is short for times of very high or very low variance and is long for times of moderate variance. This can also be explained by an intuitive understanding of clouds. Extremely volatile times will result from repeated, sharp breaks in clouds that we might expect to be less persistent in geography just as the are less persistent in time. In other words, volatility in time implies volatility in space, and cloud variability is less geographically correlated during highly volatile times than during less volatile times. For the very low volatility states the variability isn't necessarily driven by clouds, resulting in very low correlation (either through fast decay or through a small  $a$  parameter).

**4.2.1.5 Validation by Simulation** To validate our model, we simulate a distribution of the aggregate requirements from our fleet of systems,  $\sum_{i=1}^n Y_i(t)$  for all times. The covariance matrix for each observation is generated using two methods. 1) we use the most likely volatility state for each system at

Table 3: Parameters for the exponential autocovariance function of AGC requirements, dependent on the volatility state of either system.

M = 5										
S#	a					$\tau$				
	1	2	3	4	5	1	2	3	4	5
1	0.41	0.18	0.05	0.02	0.00	0.21	0.40	0.53	0.21	0.10
2	-	0.21	0.14	0.04	0.01	-	1.54	2.05	1.76	0.10
3	-	-	0.53	0.35	0.10	-	-	1.25	1.18	0.56
4	-	-	-	0.66	0.39	-	-	-	0.97	0.68
5	-	-	-	-	0.61	-	-	-	-	0.38

Table 4: Parameters for the exponential autocovariance function of load following requirements, dependent on the volatility state of either system.

M = 5										
S#	a					$\tau$				
	1	2	3	4	5	1	2	3	4	5
1	0.20	0.17	0.07	0.06	0.00	12.76	12.94	6.94	0.61	0.10
2	-	0.23	0.17	0.06	0.00	-	8.45	5.98	2.46	0.10
3	-	-	0.49	0.38	0.18	-	-	6.51	8.73	7.58
4	-	-	-	0.61	0.51	-	-	-	7.84	5.88
5	-	-	-	-	0.74	-	-	-	-	5.42

each time, as estimated during EM as  $v_{i,t} = \max_n(\gamma_{i,n,t})$ . 2) we simulate 40 sets of volatility states conditional on inputs from knots each system at each time. In the second method we simulate volatility states using the stationary probabilities of each transition matrix  $\mathbf{A}^{(k)}$ , meaning that they are conditional on  $\vec{W}_t$ , and not on the previous time-steps volatility state  $\vec{v}_t$ . This simplification is valid because the mixing rates for our process are much shorter than the study period of one month.

The first method is impossible for systems for which we don't have 1-minute generation data, we use it only as a baseline. The second method is possible, but adds some complexity to the simulation process because there are 40 sampled sets of volatility states and thus 40 sampled variances. We denote the results of each sample with a superscript, i.e.,  $\vec{v}_t^{(s)}$  and  $\Sigma_Y^{(s)}$ , where  $s \in \{1, 2, \dots, 40\}$ . We model the distribution at each time as a gaussian mixture with 40 equally likely components, one for each sampled variance.

The output of our simulation is a time series of variances that result from summing all elements of the simulated  $\Sigma_Y$  at each time point, shown for the known volatility state simulation of load following requirements in Figure 5; where the variance at each time point is used to find a separate 95% confidence interval predicting the requirement at that time. When high volatility states are simulated, the confidence interval is wider than when low volatility states are simulated.

Figures 6 and 7 show quantile-quantile (QQ) plots comparing normalized predicted quantiles to normalized empirical quantiles for the model data and test data of each metric. QQ plots are equivalent to a direct comparison of an empirical cumulative density function (CDF) to a theoretical CDF, if the two distributions are equal then their quantiles will also be equal and points in the QQ plot will lie along the  $y = x$  line. Because quantiles are normalized, the probabilities of each value can be inferred using the standard normal (e.g.  $F(-2) = 0.0228$ ). The benefit of using QQ plots instead of overlapping CDFs is that disparities over the entire range of the distribution are readily apparent in a QQ plot. This is particularly useful for viewing disparities in the distribution tails, which are difficult to show with overlapping CDFs due to their long range and fast decay.

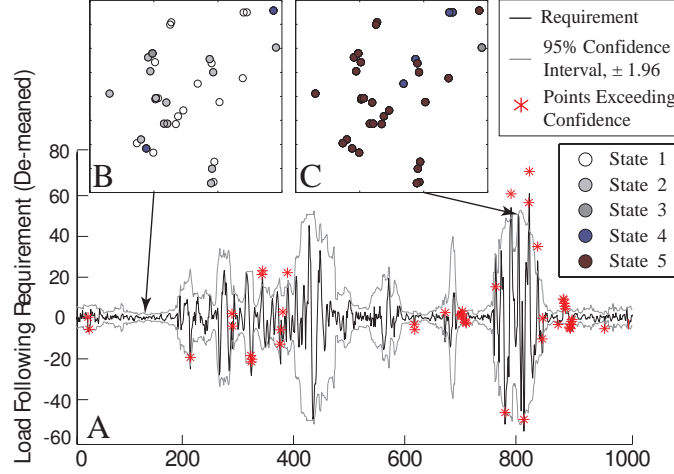


Figure 5: Output of the model: Panel A shows a time series of observed, demeaned load following requirements for the Central Valley fleet, bounded by a 95% confidence interval for each observation, points exceeding this interval are highlighted by red stars. Panels B and C show the locations of systems in Los Angeles where the color of system represents the hidden volatility state. Panel A shows volatility states at observation point 150 when total variance is low. Panel B shows volatility states at observation point 800 when the variance is high.

Columns 1 and 2 of plots in Figures 6 and 7 represent the different simulation scenarios: 1) where the most likely hidden volatility state for each system at each time is known, 2) where the model simulates independent volatility states for each system at each time conditional on the input state,  $k$ , by using the stationary probabilities of each transition matrix,  $\mathbf{A}^{(k)}$ .

As shown, the model predicts distributions almost perfectly in the baseline scenario, where the volatility state of each system at each time is the most likely estimated by EM. This result is extremely promising, it implies that if we can effectively model the volatility states of a theoretical network of PV systems then we can correctly characterize the distribution of fast time-scale impacts of distributed PV. Model 2 shows that the empirical tails are “heavy” compared to the simulated distribution, this implies that the simulation under-predicts the extreme events in the distribution. Because we see that the baseline scenario is correct, this deviation must result from ineffective prediction of volatility states.

**4.2.1.6 Conservative Assumptions for Autocorrelation of Volatility States** Our inability to effectively model volatility states prevents us from being able to correctly predict the distribution of short time-scale impacts. We believe that this disparity results from geographic autocorrelation of volatility states, meaning that systems located closely together are more likely to exhibit the same volatility state than systems that are far apart. It is possible to model such auto-correlation, but the problem is non-trivial for discrete states as most geographic autocorrelation models use continuous gaussian surfaces. In this section we develop two conservative assumptions for modeling this autocorrelation, we hope that these may provide an upper bound for extreme events in the distribution.

**4.2.1.6.1 Maximum volatility state in a neighborhood** We estimate  $\vec{v}_t$  in two stages. In the first stage we simulate volatility states exactly as defined in the model (with no geographic autocorrelation) and call this vector  $\vec{v}_t^*$ . In the second stage, we update each system’s volatility state to be the maximum element

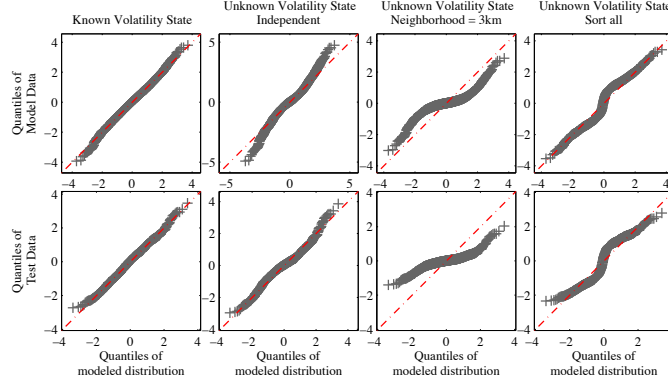


Figure 6: QQ Plots of empirical distributions versus simulated distributions for AGC requirements, results shown separately for model data and text data. Simulations of volatility states are run using four different methods 1) pre-knowledge of volatility state (baseline), 2) Independent given  $W_{t,k}$ , 3) Maximum simulated in a 3km neighborhood, and 4) 40, sorted, independent simulations.

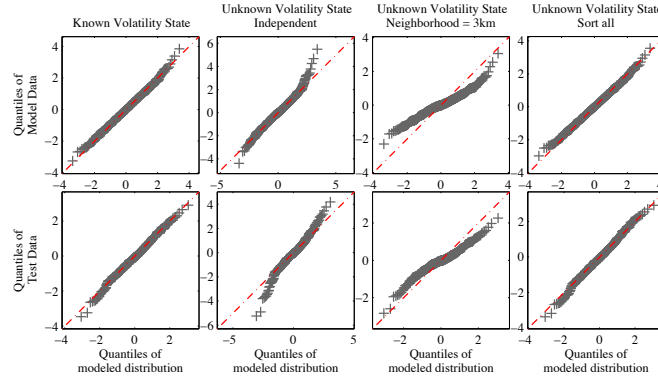


Figure 7: QQ Plots of empirical distributions versus simulated distributions for load following requirements, results shown separately for model data and text data. Simulations of volatility states are run using four different methods 1) pre-knowledge of volatility state (baseline), 2) Independent given  $W_{t,k}$ , 3) Maximum simulated in a 3km neighborhood, and 4) 40, sorted, independent simulations.

of  $v_t^*$  simulated within a neighborhood, shown in Equation 18. The neighborhood of a system  $j$  is denoted  $\mathcal{N}(j)$  and can be defined in two ways, 1, all systems within a specified distance of system  $j$ , or 2, a specified number of nearest neighbors to system  $j$ . The neighborhood of  $j$  is always inclusive of  $j$ .

$$v_{t,j} = \max_{i \in \mathcal{N}(j)} (v_{t,i}^*) \quad (18)$$

The third column of Figures 6 and 7 shows a QQ plot of the distribution simulated from this worst case assumption (using a neighborhood of 3km) against the empirical data. As shown the tails of the empirical distribution are definitively light compared to the modeled distribution, meaning that this simulation may provide us with an upper bound for predicting extreme events in the distribution.



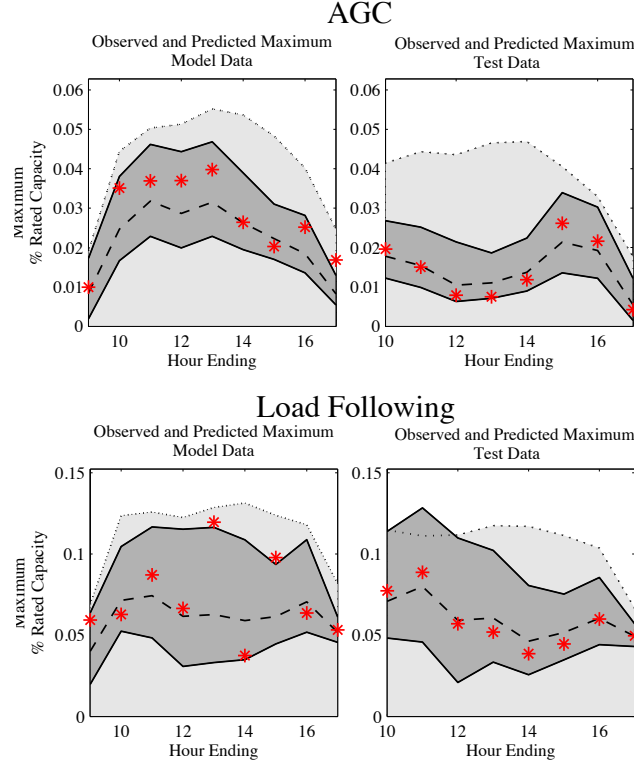


Figure 8: Predicted distributions of maxima observed at each hour ending in the test data and the model data. The dark grey boundaries signify a 95% confidence interval found using the baseline scenario where the most likely volatility states predicted by EM is used, the light grey boundary signifies the 95<sup>th</sup> percentile of the predicted distribution of maxima using the sorting based worst case assumption. Red stars signify observed maxima.

**4.2.1.6.2 Sorting sampled volatility states** Our second method for modeling geographic auto-correlation is to sort the 40 sets of volatility states simulated for each system at each time from highest to lowest variance. As a result, all systems are in their greatest sampled volatility state during the 1<sup>st</sup> sample, all systems are in their second highest state during the 2<sup>nd</sup> sample, and so on. This is an extreme model of geographic autocorrelation where we assume systems to be as autocorrelated as possible regardless of how far they are from one another.

Column 4 in Figures 6 and 7 show QQ plots for the distribution simulated with sorting versus the empirical data for AGC and load following respectively. For load following, the empirical data appears to fit the simulated distribution just as well as in the baseline scenario where volatility states are known (column 1). This signifies that significant autocorrelation of volatility states within the 256km<sup>2</sup> region may be appropriate for load following. For AGC, the empirical data appears heavy toward the mean of the distribution but then becomes lighter towards the tails; points in the plot approach the  $y = x$  line without crossing it. This signifies that the trend at the center of the distribution is to under predict empirical results, while the trend at the tails is to over predict. This is expected because volatility states are too correlated both for the high variance and the low variance values, in an ideal model they will be correlated in regions smaller than 256km<sup>2</sup>

**4.2.1.7 Predicting maximum events** Figure 8 shows predicted 95% confidence intervals for the maximum measured impact during each hour in the model data and the test data. The dark grey boundaries are those calculated where the volatility states are known, again used as a baseline. The light grey boundary is the 95<sup>th</sup> percentile of the predicted distribution using the sorting based worst-case autocorrelation method described in Section 4.2.1.6.2. As shown, the bounds defined by the ideal model are good predictors for the maximum or minimum observation over a time period. Measurements fall outside of the dark grey boundaries a total of 3 times out of 34 observations (a realization this far from an expected value of 1.7 occurs with 41% probability). Measurements are never outside of the worst-case boundary. While this is not proof that the sorting assumption always gives an upper bound of the distribution tail, it is a promising result that sorting volatility states gives a tight yet effective upper bound for worst case fluctuations.

**4.2.1.8 Discussion** Our model is effective at bounding the worst likely AGC and load following requirements of a net PV signal from an arrangement of systems. The true shape of the distribution remains elusive as we are not able to effectively predict geographic auto-correlation among volatility states. We are able to bound the tail of the distribution shape by using one of our two conservative assumptions for this autocorrelation: (1) each system takes the maximum volatility state observed in a neighborhood, or (2) sort all of the samples of simulated volatility states, an extreme form of auto-correlation where all systems are as correlated as possible.

The model can be used to predict AGC and load following requirements, under perfect forecast, for any study area for which there is consistent 15-min resolution PV generation available for input. If 15-min resolution PV generation is not available for the desired area, the same model can also be fit and validated to different, more widely available input data, e.g., satellite imagery or ground weather station observations.

Despite this achievement, our model is still very limited because it does not account for other components of net load (wind generation or electricity consumption), nor does it account for forecast error in the hour-ahead or the real-time market which will be compensated for by load following and AGC respectively. In the future, persistence of a clearness index may be used as a baseline forecast from which to measure error, such as the model presented in Ibanez et al. [21]. A similar model of geographic autocorrelation and stochastic volatility can be used to predict the forecast error in a 5-min persistence forecast. Such a model will be very interesting as it will be able to quantify the benefits of geographic diversity to PV forecasts. 60-min persistence forecast error may be estimated without a model as complicated as this one by leveraging the 15-min generation data available through California. This can be used in conjunction with the above models to include forecast errors in our predictions.

Other components of net-load can be included by (1) adjusting the mean of the AGC and load following requirements to include those simulated for wind and/or consumption or (2) including similar statistical models of consumption and/or wind.

## 4.2.2 Predicting the Regulation and Load Following Requirements of Distributed Photovoltaics in California

**4.2.2.1 Introduction** Grid-connected solar photovoltaic (PV) systems are rapidly increasing in number. Over 200,000 PV systems are installed in the US, amounting to over 3.5 GW of total rated capacity. Most of these installations are concentrated in cities in the West and the Desert Southwest. For example, there are over 1.5GW in California alone; 550MW of these are in the counties of Los Angeles, Riverside, San Diego, and Santa Clara. California's goals for the next decade are even more ambitious: Combined initiatives include subsidies for 3 GW of distributed PV by 2016 [2], and future goals aim to increase this amount four-fold to 12 GW of distributed generation (mostly PV) by 2022 [3].

This rise in distributed PV coincides with concerns that intermittent generation from wind and solar will compromise the stability of the electricity delivery system and lead to increased costs for maintaining reliability. Specifically, additional, uncontrolled variability from wind and solar may require additional, fast-responding generators that can take up to *10 years?* to go through permitting and construction. As a result, many organizations concerned with grid planning have attempted to quantify the requirements of future variability. Some of the most frequently cited integration studies are the 20% (and the forthcoming 33%) renewables integration study by the California Independent System Operator (CAISO) [7], and the National Renewable Energy Laboratory's Western (and Eastern) Wind and Solar Integration Studies [4]. For analyses related to wind generation, there is the NREL dataset of 10min resolution predicted wind generation data throughout the country from 2004 to 2006. This data is often considered sufficient for wind because it is less intermittent at short time scales. However, though there exist similar datasets for PV, they are often deemed insufficient. PV is known to fluctuate dramatically at 1-minute resolution and even 1-second resolution monitoring [22].

Both studies rely on a limited amount of 1-minute resolution generation data from intermittent solar PV generation for their analysis. In the case of the CAISO 33% study, sub-hourly PV generation data is repeated multiple times in order to fill gaps during the study period, for the NREL study, a statistical analysis relating ramp rates in load, wind and solar generation is completed however only used to a very limited extent to predict the actual sub-hourly requirements from intermittent generation.

Using statistics to assess the amount of fast-responding resources that distributed PV will require of power systems remains difficult. Particularly because power systems planners are concerned with high-impact, low-probability events, thus any statistical model will need to effectively predict extreme events instead of common events. As described in the Chapter 4.2.1, current statistical models for predicting variability from PV are capable of effectively predicting the variance of of a distribution of fluctuations from a fleet of solar PV systems. However without a specific distribution shape, the mean and variance alone are not sufficient for predicting extremes. Robust statistical models of variability from renewables are needed to accurately assess the future needs of the power system without relying on expensive, conservative assumptions.

In this Part, we apply the model from Chapter 4.2.1 to predict extreme events from solar PV system located through California during the study period of the 17<sup>th</sup> of November 2011 to the 17<sup>th</sup> of November 2012. Specifically, we predict the operational requirements for Automatic Generation Control (AGC) and load following to follow the variable generation of 6GW and 12GW of solar PV in California.

This study is limited for two important reasons. First, our model only predicts the AGC and load following needs to follow the negative load signal from solar PV generation in California, in actuality total net load also includes electricity consumption and negative load from other renewables. While this is not ideal,

this section is first a proof of concept for this modeling strategy, which can later be used in conjunction with predicted net load from wind and consumption. Also, this Part compares different scenarios for the future of PV in California and can still answer important policy questions regarding where to site PV systems and how to balance them. We compare the requirements for three possible arrangements of PV systems, representing different policy future for siting and incentivising PV:

1. **fully distributed** all PV is installed in 5kW systems uniformly distributed throughout our study area.
2. **fully community-scale** all PV is installed into 3000 community-scale systems (either 2MW, or 4 MW systems) throughout populated areas of California.
3. **fully centralized** all PV is installed in 200MW systems in areas with the largest available resource.

We also compare four possible sizes of balancing areas, these regions demonstrate the various balancing needs for solar PV depending on which entity is responsible for balancing variability from PV:

1. 4 square kilometer community balancing,
2. balancing by county, and
3. balancing all of California.

Second, our current model is limited because it predicts AGC and load following requirements under perfect forecast.

**4.2.2.2 Simulation Model** We simulate the AGC and load following requirements using the model described, fit and validated in Chapter 4.2.1, *to be submit to IEEE Transactions on Power Systems*. To reduce computational requirements for simulating over 1 million PV systems, we adjust the simulation method to run in two stages: the first stage simulates variability at a fine geographic resolution and is generally applicable across time and location, the second stage simulates a coarse resolution that specific to locations in California during our study period.

The coarse resolution consists of a 2km by 2km partial grid of California, shown in Figure 9. We only include grid cells for which we have complete data during the study period (described in Section 4.2.2.2.2), amounting to 2407 grid cells covering 9628 km<sup>2</sup>. While this is only a small portion of California, our coverage is more than half that of California's urban areas, where over 80% of Californian's live, these areas are also shown in Figure 9 [23].

Each grid cell is modeled to contain one of five possible, characteristic arrangements of PV systems, corresponding to one or more of our siting scenarios. Table 4.2.2.2 describes how the arrangements of systems within individual cells are used to model each scenario. For the distributed and community scenarios, all 2407 available grid cells contain PV systems. For the centralized scenario, only enough grid cells are filled to reach the desired total capacity of PV in California: 15 for the 6GW scenario and 30 for the 12GW scenario.

Figure 10 displays the characteristic arrangements of PV systems used in our model. Because our model is fit and validated on small, 2kW to 12kW PV systems, modeling large PV plants as one system would likely ignore within-plant geographic diversity. Community and Centralized systems are modeled as grids of 5kW PV systems with 10 meter spacing between. A 5kW nameplate rated PV system with a 20% efficiency covers about 25m<sup>2</sup>, our 10m × 10m spacing covers 100m<sup>2</sup>, representing a ground cover ratio of 0.25.

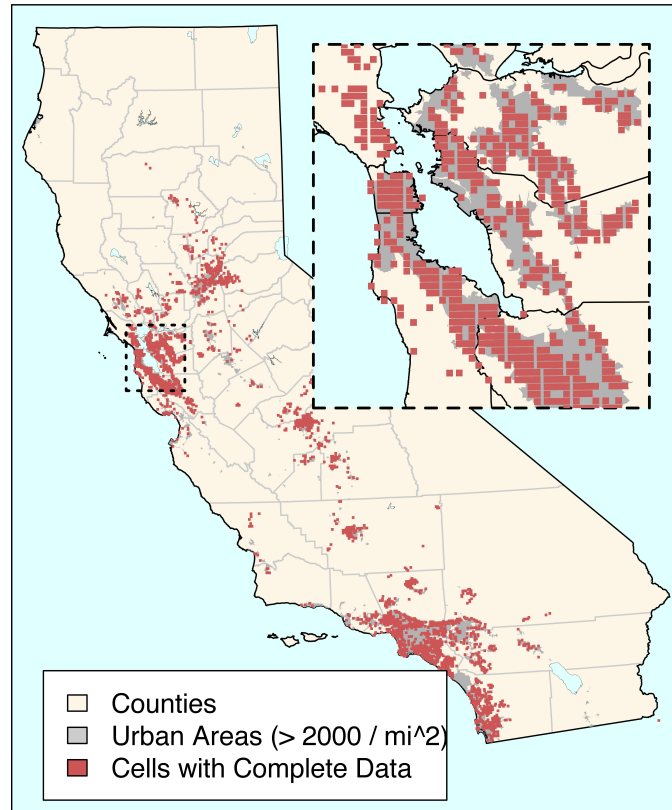


Figure 9: Map of grid cells for which there exists consistent 15-min resolution generation data from Sept. 15th 2011 to Sept. 15th 2012; also California county boundaries and urban areas with greater than 2000 people per square kilometer

This is a reasonable value for utility-scale PV systems [24].

PV systems are dispersed evenly throughout all available grid cells in the Community and Distributed scenarios. For the Centralized scenario, we include as much PV as possible within a grid cell (200MW) and then only fill the cells with the greatest annual PV resource, as reported by NREL [25], and shown in Figure 11.

**4.2.2.2.1 Balancing Areas** We test four possible of sets balancing areas used for managing variability: (1) each 4 square kilometer grid cell is balanced separately, (2) each county in California is balanced separately (3) the entirety of California is one balancing area.

We report the total balancing requirements for all systems in California for each set of balancing areas. I.e., we find the total power capacity required to meet the worst likely fluctuation in each balancing area, and call this the capacity requirement for that area; we then add all the requirements in all areas in a set to find the total capacity requirement for California.

We expect that the required capacity will decrease as the size of the balancing area increases, because the maximum requirements in the smaller balancing areas are not likely to occur simultaneously. Thus we don't have to prepare for all of them to occur at once.

Table 5: Description of scenarios

Scenario	Nameplate Capacity	# of cells	Arrangement of systems within grid cells
Distributed	6 GW	2407	500, randomly located, 5kW PV systems
Distributed	12 GW	2407	1000, randomly located, 5kW PV systems
Community	6 GW	2407	1, 2.5 MW PV system
Community	12 GW	2407	1, 5.0 MW PV system
Centralized	6 GW	30	1, 200 MW PV system
Centralized	12 GW	60	1, 200 MW PV system

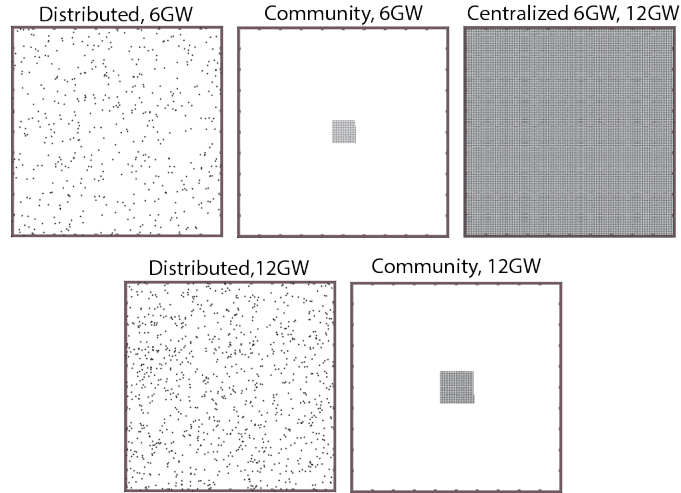


Figure 10: Characteristic arrangements of PV systems used for each scenario. Arrangements are as described in Table 4.2.2.2

**4.2.2.2.2 Input Data** All systems in a grid cell are dependent on the same time series of input heuristics, defined in Section 4.2.1.2.3. For each grid cell, we attempt to assemble consistent 15-min resolution generation data for our study period of the 17th of September 2011 to the 17th of September 2012. This time-series is used to calculate the input heuristic for the simulation model.

Instantaneous voltage and current at 15 minute resolution are available for over 6000 systems in California from January 2011 to late September 2012. Monitored PV systems are owned and maintained by SolarCity., the data is provided to us under the terms of a nondisclosure agreement. Data were collected at the inverters for each system and sent to servers via ZigBee devices, communication limitations resulted in large gaps in data collection.

All consistent data from the closest system to each cell's centroid is included in that cell's time-series. Gaps in the time series are filled using data from the next closest system. Data is evaluated for consistency on a daily basis, thus for each day in a grid cell's time series, all data is sourced from the same system. The gap filling process is repeated for systems progressively further from the centroid until a complete set of data is obtained for the cell, or until there are no more systems left within a 5km radius. Only grid cells that obtain greater than 80% of their data from systems within the cell are used for simulation, amounting to the 2407 cells shown in Figure 9.

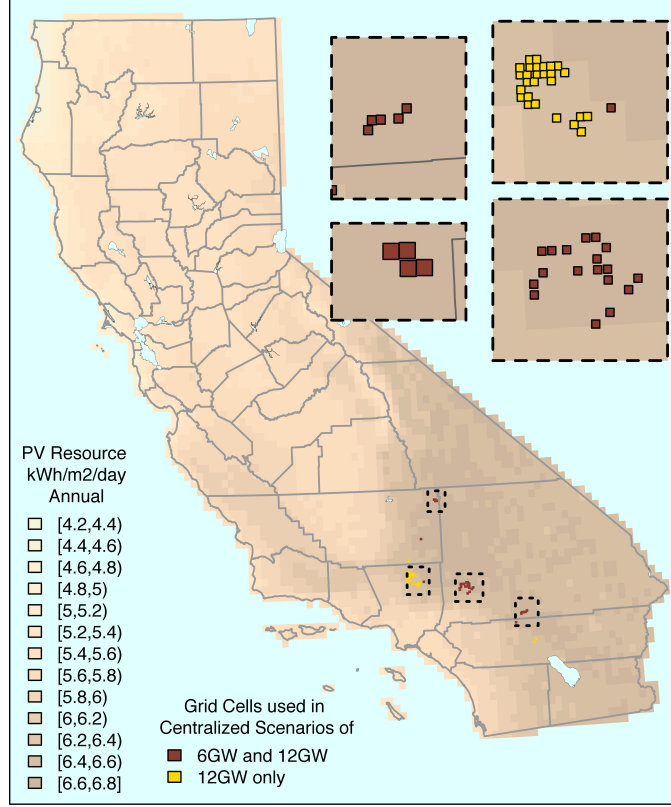


Figure 11: Map of grid cells used for the centralized scenarios. Available cells with the greatest annual PV resource are chosen.

**4.2.2.2.3 Stage 1: Fine Resolution Modeling** Fine resolution simulations are used to sample covariances between aggregate generation from PV systems in pairs of grid cells. This covariance is independent of time and space when normalized by the clear sky signal, thus one set of samples is applied throughout California during the entire study period. Each pair of grid cells is defined by

1.  $a = \{a_1, a_2\}$ , The pair of arrangements of PV systems contained within each cell.
2.  $w = \{W_1, W_2\}$ , The pair of input states for each cell.
3.  $d$ , The distance separating the pair of cells. We simulate distances of 0km and 2 to 20km at intervals of 0.5km.

There are 5 possible arrangements of PV systems within a cell, leading to  $\frac{5 \cdot 6}{2} = 15$  possible combinations of arrangements. For each pair, there are  $5^2 = 25$  possible combinations of input states. Including the 38 simulated distances, there are a total of  $15 \cdot 25 \cdot 38 = 14250$  characteristic pairs of grid cells.

For each PV system in a characteristic pair of cells, we simulate 200 samples of volatility states assuming geographic independence. The conservative assumption of sorting volatility states, described in Section 4.2.1.6.2, is applied so that all of the greatest variance volatility states for each system occur during the same sample. This is a model of extreme geographic auto-correlation of volatility states that ensures the tails of the final distribution will be heavy compared to empirical distribution tails.

The result of the each sample is a large,  $N \times N$  covariance matrix,  $\Sigma^{(a,w,d,s)}$ , with a row and column for each system in either grid cell; here  $N$  signifies the total number of systems. The  $s$  subscript signifies the sample number, where 1 contains the most volatile simulated states and 200 contains the least.

The covariance matrix relating the aggregate generation from all systems in each cell of the characteristic pair is found using Equation 19; where  $H$  is an  $N \times 2$  matrix of zeroes and ones where each column identifies the systems in one grid cell.

$$\hat{\Sigma}^{(a,w,d,s)} = H' \Sigma^{(a,w,d,s)} H \quad (19)$$

$\hat{\Sigma}^{(a,w,d,s)}$  is a  $2 \times 2$  symmetric matrix, the diagonals are the variances of aggregate generation within each grid cell, and the off-diagonals are the covariance between aggregate generation from each grid cell in the pair. Only the covariance (on the off-diagonals) of each sample is stored and used in Stage 2, referred to as  $\sigma_{a,w,d,s}$ .

**4.2.2.2.4 Stage 2: Coarse Resolution Modeling** In the coarse resolution model, we use the results of the fine resolution model in Section 4.2.2.2.3, and the input data described in Section 4.2.2.2.2, to model the requirements for following specific arrangements of PV systems in California during the study period. At each time-step, we model aggregate generation from all systems within each grid cell rather than generation from each system individually. This reduces the number of units of analysis from  $> 10^6$  systems to only 2407 grid cells; thus the number of elements of each covariance matrix is reduced from  $> 10^{12}$  to  $2407^2$ .

For each 15-min interval,  $\tau$ , during the study period, we know the input state and the location of each grid cell; we specify an arrangement of PV systems within the cell depending on the scenario modeled, shown in Table 4.2.2.2. With this information, we used the sampled covariances from Stage 1 to populate a  $2407 \times 2407$  covariance matrix for the normalized requirement signals from aggregate generation each grid cell in California, denoted  $\Sigma_{grid}^{(\tau,n)}$ . We generate 5 covariances matrixes for each 15-min period. For each covariance matrix, one sample number,  $s$ , is randomly chosen and is used to populate all elements of the matrix. Using the same  $s$  for all elements of  $\Sigma_{grid}^{(\tau,n)}$  is effectively an extreme case of geographic autocorrelation of volatility states, similar to simulating the volatility states for all of the systems in California and then sorting all of them together,

Distances between grid cells in California are rounded to the nearest half kilometer to correspond to distances simulated in Stage 1, and distances of greater 22.25 km are assumed to result a covariance of zero.

A clear sky signal is simulated for each grid cell using a solar-earth geometry model described in Section 4.2.1.2.2 For distributed arrangements, this clear sky signal is simulated assuming that the distribution of system geometries within the cell matches the distribution from system geometries in our metadata, described in Section 4.2.1.2. All systems are assumed to have a derating factor of 0.77. For centralized and community scenarios, all systems are modeled to face due south with a tilt equal to their latitude. The AGC and load following requirements for the clear sky generation from each grid cell are calculated, denoted  $\vec{y}_{CL}(t) \in \mathbb{R}^{2407}$ . For AGC there is a requirement every minute, thus there are 15 realizations per 15-min interval. For load following there are 3 realizations per 15 minute interval.

To “un-normalize” the covariance matrices at each time, the diagonal matrix  $M^{(t)}$  is constructed using hourly maxima of the clear sky signal, shown in Equation 20; where  $CL_i(t)$  is the clear sky signal in grid



Table 6: Total capacity requirements predicted for each scenario and each balancing area. Total capacity requirements are the 95<sup>th</sup> percentile of predicted maximum requirements during the study period of one year

Balancing Area	6 GW of Nameplate Capacity						12 GW of Nameplate Capacity					
	Distributed		Community		Centralized		Distributed		Community		Centralized	
	Down	Up	Down	Up	Down	Up	Down	Up	Down	Up	Down	Up
<b>AGC</b>												
4km <sup>2</sup> Region	-829	829	-2223	2223	-800	801	-1613	1614	-4153	4154	-1593	1594
County	-91	92	-224	224	-222	223	-179	180	-418	419	-384	386
State	-17	17	-40	41	-142	142	-33	34	-75	76	-203	204
<b>Load Following</b>												
4km <sup>2</sup> Region	-2416	2437	-2803	2829	-2532	2559	-4853	4898	-5602	5654	-5052	5101
County	-811	817	-878	891	-1293	1315	-1613	1629	-1764	1790	-2593	2629
State	-427	398	-549	497	-949	966	-859	790	-1094	998	-1619	1634

cell  $i$ .

$$M_{i,i}^{(t)} = \max_{t \in \text{hour}} CL_i(t) \quad (20)$$

Equation 21 shows a gaussian mixture distribution that models the AGC or load following requirement for generation in each grid cell at each time, denoted  $\vec{y}_{grid}(t)$ . In Equation 21,  $\tau$  is the 15-min interval containing  $t$ .

$$\vec{y}_{grid}(t) \sim \frac{1}{5} \sum_{n=1}^5 MVN \left( \vec{y}_{CL}(t), M^{(t)} \Sigma_{grid}^{(\tau,n)} M^{(t)} \right) \quad (21)$$

Equation 22 shows the distribution of a requirement signal for aggregate generation within a balancing area; where  $y_a$  is the AGC or load following requirement to follow aggregate generation in balancing area  $a$ ,  $\vec{u}_a$  is a vector with ones corresponding to grid cells within the balancing area and zeros otherwise, and  $\Phi$  is the CDF of a univariate normal.

$$p\{y_a(t) \leq x\} = p\{\vec{u}_a' \vec{y}_{grid}(t) < x\} = \dots \quad (22)$$

$$\frac{1}{N} \sum_{n=1}^N \Phi \left( \vec{u}_a' \vec{y}_{CL}(t), \vec{u}_a' M^{(t)} \Sigma_{grid}^{(\tau,n)} M^{(t)} \vec{u}_a \right)$$

The CDF of the maximum requirement during a time period is the product of the CDFs at each interval during the time period, shown in Equation 23 for time period  $T$ .

$$p \left\{ \max_{t \in T} [y_a(t)] \leq x \right\} = \prod_{t \in T} p\{y_a(t) \leq x\} \quad (23)$$

We are concerned with extreme events, thus we report the 95<sup>th</sup> percentile of likely maximum requirements during a time period as the total power capacity required for balancing.

**4.2.2.3 Results: Total Capacity Requirements for California** Table 6 shows the projected capacity requirements for each scenario and for each balancing area size, results are also displayed separately for

load following up and down. Each capacity requirement is the 95<sup>th</sup> percentile of predicted requirements during the study period of one year.

By far the most dramatic reductions in balancing need are achieved by increasing the geographic size of the balancing area. In every scenario, increasing the balancing area from 4km<sup>2</sup> regions to Statewide balancing decreases the total balancing capacity required by at least 80% for AGC requirements and by at least 60% for load following. This means that requiring community solar or distributed solar to firm signals at the system would require at least a 2 to 5 fold increase in the total amount of storage capacity needed for California.

While not as dramatic as the effect of balancing area, the effect of geographic diversity of PV systems in California also presents an interesting story. For AGC, the state-wide balancing needs from solar PV under perfect forecast are minimal, peaking at 200MW for 12GW of installed PV capacity! However the "perfect forecast" here is a significant qualification, errors in forecasts that are made to dispatch load following translate into extra AGC requirements. Since these forecast take place on a 5-min time-scale, we expect them to exhibit autocorrelation similar to load following here, but not to be as large in magnitude since it is following a mean event instead of extreme events.

System dispersion within a grid cell has a dramatic effect on AGC requirements. As shown in Table 3, the autocorrelation functions for the AGC models have very short decay distances, typically less than 1 km. Thus dispersing systems through a 2km×2km grid cell results in a dramatic reduction in AGC requirements, shown in the comparison of the distributed scenario to community scenario or the centralized scenario.

Load following does not display this effect. As shown in Table 4, the autocorrelation functions for load following have decay distances of longer than 2 km. Thus systems are similarly correlated regardless of their relative locations within a grid cell. This is shown by the relatively small difference between the distributed and community scenarios for load following. However, the dispersion of grid cells with PV in California did have a large effect on the load following balancing need, shown by the difference between the centralized scenarios and the distributed scenarios. In the centralized scenarios, PV is located only in a few, closely located grid cells which result in a significant amount more autocorrelation than when located in 2407 grid cells dispersed throughout California.

**4.2.2.4 Results: Hourly Capacity Requirements for California** Figures 12 and 13 show the predicted total capacity requirements on an hourly basis. In this case, we find the 95<sup>th</sup> percentile of predicted maximum requirements within an hour ending over the entire year. Bars extending above the x-axis represent regulation-up, bars below the x-axis represent regulation-down. These values will, in general, be lower than those reported in Table 6 because they are the maxima of less realizations of a random variable. This reporting method is similar to that in Makarov et. al and is useful because it points out when the capacity is needed. Some dispatchable generation (e.g., hydro) is only available at certain times of day. Also, more dispatchable generation is available during off-peak times than on-peak, because less generators are needed to meet predicted base load.

Figures 12 and 13 show three overlapping columns for each hour ending and each scenario: the tallest, light-grey bar represents the total capacity needed if balancing areas are 4km<sup>2</sup>, the light blue bar represents the county level balancing scenario, and the dark blue bar represents the statewide balancing scenario. Also shown is a black line representing the balancing need under clear sky conditions, this represents a baseline, the minimum possible balancing need achievable by increasing balancing areas. For

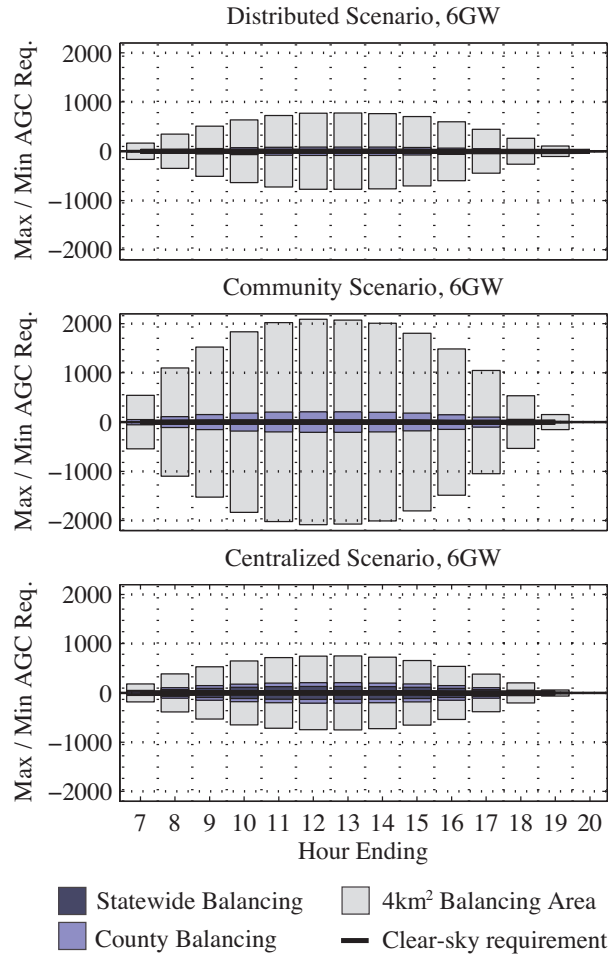


Figure 12: Predicted maximum AGC requirements during each hour ending for the distributed scenario (top), community scenario (middle), and centralized scenario. Each scenario reflects 6GW of total nameplate capacity of PV in California. Colors represent different sized balancing areas, the black line represents the requirement from the clear sky signal (in absence of clouds).

AGC, this line is essentially 0, there is very little need for AGC to follow solar PV generation under clear sky conditions. For load following, this line exhibits peaks in the morning and evening shoulders, when load following is needed to meet the dramatic ramp up and down of solar electricity with the rising and setting sun.

As with the total requirements in Section 4.2.2.3, dramatic reductions are seen by increasing the size of the balancing area. For the Distributed and Community scenarios, state-wide level balancing decreases capacity requirements almost to those required by the clear sky signal. The centralized scenario does not hit this limit, at mid-day hours the load following capacity required to follow centralized PV is much larger than the clear sky limit. Thus, for load following, concentrating PV arrays in large groups, such as the ones shown in Figure 11 leads to increased need for load following beyond that required to follow the clear sky signal, where distributing PV throughout California allows state-wide system balancers to need to account for cloud variability.

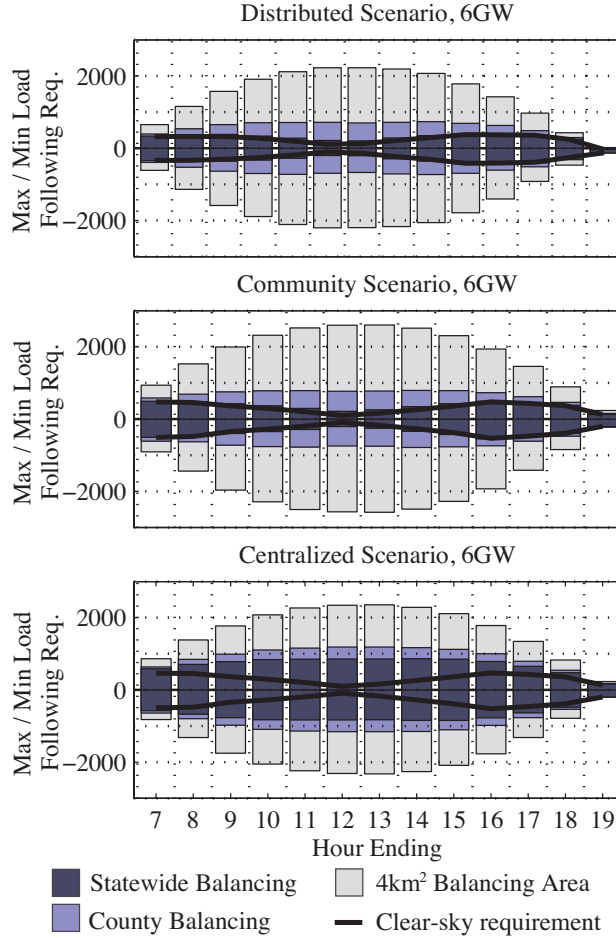


Figure 13: Predicted maximum load following requirements during each hour ending for the distributed scenario (top), community scenario (middle), and centralized scenario. Each scenario reflects 6GW of total nameplate capacity of PV in California. Colors represent different sized balancing areas, the black line represents the requirement from the clear sky signal (in absence of clouds).

**4.2.2.5 Discussion** When considering state-wide system balancing of net load, these results are insufficient: they do not address the uncertainty of the signal and they do not include other important components of net load. Despite these qualifications, there are still important lessons that we learn from this simulation.

There are major gains from balancing in wide areas versus small ones, these should be considered when writing policy on energy storage and distribute PV. Current trends in policy incentives firming variable generation signals from PV directly at the system, even small systems of 2-10 kW capacity. Firming on this level may have benefits for the distribution system, but it should be noted that these small areas will lead to dramatic increases for the total required power capacity for balancing in California. This increase results from redundant use of balancing resources, some will increase while others are decreasing. Our model estimates the power capacity needs for balancing in California, not the energy capacity needs, however the reduction in power capacity requirements here are likely also indicative of a reduction in energy capacity requirements. As the load following and AGC requirement signals are aggregated over wider areas, it is likely that it will become less temporally autocorrelated. Thus load following up or down events

are likely to be shorter in duration and requires less energy capacity related to the power capacity.

Centralized PV exhibits larger AGC and load following requirements than Distributed or Community PV systems. Distributed and Community PV has the ability to reach extremely high amounts of installed capacity without adding any balancing requirement due to un-predictable could cover. This effect is, however, not as large as the effect of sizing the balancing area.

This study is limited by its assumptions of (1) perfect forecast for solar PV and (2) accounting for only the solar PV component of net load, however, future advancement can include these effects relatively easily. Persistence forecasts of a clearness index can be used to model forecast error in the hour ahead market and the real time market, such as the one presented in [21]. Another model of geographic autocorrelation with stochastic volatility (similar or identical to the one presented in this section) can be used to predict the errors of a 5-min persistence forecast for solar PV. Such a model will be very useful as it will predict forecast error in solar PV while accounting for geographic diversity. Persistence forecast error for the hour-ahead market may be predictable the our 15-min generation without a model such as the one presented in this section.

The wind component and the consumption component of net load can be included in two ways: (1) the mean used in this model can include the AGC or load following requirements from simulated net load found in other studies, or (2) a similar model to this one can be applied to consumption and/or wind generation. A set of these models can be combined to predict the maximum requirement of the sum of each signal.

## References

- [1] N. R. E. Laboratory, "The open PV project." [Online]. Available: <https://openpv.nrel.gov/>
- [2] "The california solar initiative," Aug. 2013. [Online]. Available: <http://www.gosolarcalifornia.org/about/csi.php>
- [3] The California Governor's Office of Planning and Research, "Renewable energy in california." [Online]. Available: [http://opr.ca.gov/s\\_renewableenergy.php](http://opr.ca.gov/s_renewableenergy.php)
- [4] R. Piwko, K. Clark, L. Freeman, G. Jordan, and N. Miller, "Western wind and solar integration study: Executive summary,(WWSIS) may 2010," National Renewable Energy Laboratory (NREL), Golden, CO., Tech. Rep., 2010.
- [5] X. Bai, K. Clark, G. A. Jordan, N. W. Miller, and R. J. Piwko, "Intermittency analysis project: Appendix b impact of intermittent generation on operation of california power grid," California Energy Commission, Tech. Rep., 2007.
- [6] Y. Makarov, C. Loutan, J. Ma, and P. de Mello, "Operational impacts of wind generation on california power systems," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 1039–1050, May 2009. [Online]. Available: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4808228>
- [7] M. Rothleder, "Track i direct testimony of mark rothleder on behalf of the california independent system operator corporation," California Public Utilities Commission, 2011.
- [8] M. Lave, J. Kleissl, and E. Arias-Castro, "High-frequency irradiance fluctuations and geographic smoothing," *Solar Energy*, no. 0. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0038092X11002611>
- [9] E. Wiemken, H. G. Beyer, W. Heydenreich, and K. Kiefer, "Power characteristics of PV ensembles: experiences from the combined power production of 100 grid connected PV systems distributed over the area of germany," *Solar Energy*, vol. 70, no. 6, pp. 513–518, 2001. [Online]. Available: <http://www.sciencedirect.com/science/article/B6V50-42PBYFM-7/2/d17aa4bb4d5e3fe178818a10a583448e>
- [10] A. Murata, H. Yamaguchi, and K. Otani, "A method of estimating the output fluctuation of many photovoltaic power generation systems dispersed in a wide area," *Electrical Engineering in Japan*, vol. 166, no. 4, pp. 9–19, Mar. 2009. [Online]. Available: <http://onlinelibrary.wiley.com/doi/10.1002/eej.20723/abstract>
- [11] A. Mills and R. Wiser, "Implications of wide-area geographic diversity for short-term variability of solar power," Lawrence Berkeley National Laboratory, Tech. Rep., 2010.
- [12] T. E. Hoff and R. Perez, "Quantifying PV power output variability," *Solar Energy*, vol. 84, no. 10, pp. 1782–1793, Oct. 2010. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0038092X10002380>
- [13] M. Lave, J. Kleissl, and J. Stein, "A wavelet-based variability model (WVM) for solar PV power plants," *IEEE Transactions on Sustainable Energy*, vol. 4, no. 2, pp. 501–509, 2013.

- [14] M. Lave and J. Kleissl, "Cloud speed impact on solar variability scaling – application to the wavelet variability model," *Solar Energy*, vol. 91, pp. 11–21, May 2013. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0038092X13000406>
- [15] R. Perez, S. Kivalov, J. Schlemmer, K. Hemker Jr., D. Renné, and T. E. Hoff, "Validation of short and medium term operational solar radiation forecasts in the US," *Solar Energy*, vol. 84, no. 12, pp. 2161–2172, Dec. 2010. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0038092X10002823>
- [16] R. Perez, S. Kivalov, J. Schlemmer, K. Hemker Jr., and T. Hoff, "Parameterization of site-specific short-term irradiance variability," *Solar Energy*, vol. 85, no. 7, pp. 1343–1353, Jul. 2011. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0038092X11000995>
- [17] M. Hummon, E. Ibanez, G. Brinkman, and D. Lew, "Sub-hour solar data for power system modeling from static spatial variability analysis," National Renewable Energy Laboratory, Tech. Rep. NREL/CP-6A20-56204, 2012. [Online]. Available: <http://www.nrel.gov/docs/fy13osti/56204.pdf>
- [18] M. J. Reno and J. Stein, "Using cloud classification to model solar variability," Sandia National Laboratories, Tech. Rep., 2013. [Online]. Available: [http://energy.sandia.gov/wp/wp-content/gallery/uploads/SAND-2013-1692C\\_ASES-CloudCategoryVariability.pdf](http://energy.sandia.gov/wp/wp-content/gallery/uploads/SAND-2013-1692C_ASES-CloudCategoryVariability.pdf)
- [19] M. D. Tabone and D. S. Callaway, "Parameterizing fluctuations in solar photovoltaic generation using hidden markov models," in *Proceedings of the 2013 IEEE Power & Energy Society General Meeting*, Vancouver, BC, Jul. 2013.
- [20] G. M. Masters, *Renewable and Efficient Electric Power Systems*, 1st ed. Wiley-IEEE Press, Aug. 2004.
- [21] E. Ibanez, G. Brinkman, M. Hummon, and D. Lew, "A solar reserve methodology for renewable energy integration studies based on sub-hourly variability analysis," in *2nd International Workshop on Integration of Solar Power in Power Systems Proceedings, Lisbon, Portugal*, 2012. [Online]. Available: <http://www.nrel.gov/docs/fy12osti/56169.pdf>
- [22] J. Wergner, M. Lave, J. Luoma, and J. Kleissl, "Temporal downscaling of irradiance data via hidden markov models on wavelet coefficients: Application to california solar initiative data." 2012.
- [23] "State of california GeoPortal." [Online]. Available: <http://portal.gis.ca.gov/geoportal/catalog/main/home.page>
- [24] A. Goodrich, T. James, and M. Woodhouse, "Residential, commercial, and utility-scale photovoltaic (PV) system prices in the united states: Current drivers and cost-reduction opportunities," National Renewable Energy Lab, Golden, CO, Tech. Rep. NREL/TP-6A20-53347, Feb. 2012.
- [25] National Renewable Energy Laboratory, "Solar maps." [Online]. Available: <http://www.nrel.gov/gis/solar.html>

## 4.3 Optimal Rate Designs and ISO Services for Maximizing the Value of Combined PV and Storage

### Summary

This is one of three final report sections completed by a UC Berkeley team of students and faculty in the Energy and Resources Group. It covers the efforts of Tasks 4.3 and 4.5. Our specific focus is on understanding distribution system impacts and the opportunity for creating value by incorporating storage into the CAISO's dispatch process.

In Part 1 (Section 4.3.1) of this deliverable, we study the effects of deployment of high-penetration photovoltaic (PV) power on the distribution grid. The magnitude of these effects may vary greatly depending upon feeder topology, climate, PV penetration level, and other factors. We will present a simulation study of eight representative distribution feeders in three California climates at PV penetration levels up to 100%, supported by a unique database of distributed PV generation data. We find that PV penetrations up to 50% reduce system losses and feeder peak loads while having positive or negligible effects on transformer aging, voltage regulator wear, and voltage quality. At higher penetrations we observe diminishing benefits for system losses and, in some scenarios, undesirable impacts on other metrics. We present data illustrating the range of variation with feeder topology and climate. We then use our physical results and an economic model of distribution expenditures supported by data from PG&E to estimate various economic impacts of PV. We find that among the benefits we can calculate the avoided energy cost is clearly the largest (at approximately \$0.035/kWh of PV produced), with avoided distribution capacity costs being smaller but potentially significant (as much as \$0.007/kWh).

In Part 2 (Section 4.3.2), we address pricing mechanisms to improve the cost and quality of frequency regulation. We propose a regulation pricing methodology that is analogous to locational marginal prices in economic dispatch. In this way the prices are directly tied to the physics of the system and represent an improvement over the "mileage payments" currently being used to pay for regulation service performance. We also study a market design that will induce regulation providers to bid regulation services competitively. We demonstrate this feature by combining traditional frequency regulation, area control error, and the California Independent System Operator's Flexible Ramping Product in an example.

In Part 3 (Section 4.3.3), we analyze strategic behavior between non-generating resources (NGRs) providing fast regulation in reserve markets. We apply a two-stage framework in which firms first declare capacities and then engage in price bidding; the latter stage is commonly known as Bertrand-Edgeworth competition. We show the conditions under which the market has unique equilibria. By applying the model to energy storage competition, we obtain direct comparisons between two representative market formats based on capacity and energy payments. We find that energy payments may lead to slightly more predictable NGR capacity commitment and reduced regulation energy prices.



### 4.3.1 Physical and Economic Effects of Distributed PV Generation on California's Distribution System

**4.3.1.1 Introduction** As the deployment of distributed photovoltaics (PV) has accelerated in recent years, researchers and power industry professionals have increasingly attended to the impacts – both positive and negative – that such generation may have on the distribution system. Areas of concern include PV's effect on:

- System losses
- Peak load (which impacts capacity investments)
- Transformer aging
- Voltage regulator mechanical wear
- Power quality, particularly voltage magnitude
- Reverse power flow and its effect on protection systems









Prior work in this area has largely consisted of case studies that examine a selection of these issues in detail for a particular feeder. For example, studies have modeled the economic transmission and distribution benefits of siting PV at a specific substation [72] and applied sampled solar irradiance data to a simulated distribution line with PV to assess associated voltage fluctuations [77]. Because the distribution system is highly heterogeneous in terms of topology, climate and loads served, it can be difficult to draw useful generalizations from such case studies.

More recent work has begun to address this by simulating a wider variety of feeders and presenting a more complete range of results. One such study found that PV reduced line losses and did not create problematic voltage issues at penetrations up to 100% (1 kW/household) under a variety of European feeder topologies and climates, with some voltage issues and increased losses at higher penetrations [61]. Another study found that many simulated U.S. feeders could handle high penetrations of PV without voltage or overcurrent issues [45]. In both of these studies, outcomes varied significantly depending on feeder topology, PV placement, and other factors.

In this study we seek to extend this more comprehensive line of modeling by examining several metrics of PV's effect on distribution feeders typical of California. We simulate several feeder topologies under both Northern and Southern California weather conditions for one sample year in order to capture a range of effects that might be expected in different climates. We then feed the simulation results into economic models informed by financial data from Pacific Gas and Electric (PG&E) to estimate the economic value of distributed PV power in PG&E's service territory. Our work is supported by a unique database of PV generation measurements that enables us to incorporate small spatial scale geographic diversity of PV generation into our models.

**4.3.1.2 The Distribution Feeder Models** We used GridLAB-D version 2.3 (with the forward-backward sweep power flow solver) to model distribution circuits due to its integration of power flow analysis and time-varying load models, availability of representative feeder models, and open-source license. In this section we describe our preparation of the models and supporting data.

Table 1: Summary of Simulated Feeder Characteristics and Figure Legend

	Name*	Serves [67]	Peak Load (MW) [67]	Dist. Transformers	Avg House (kW) [4]	Approx Length (km)	Baseline Peak Load (MW)			Solar Profiles Available		
							Berk.	L.A.	Sac.	Berk.	L.A.	Sac.
	R1-12.47-1	mod. suburban & rural	7.15	618	4.0	5.5	5.56	5.38	7.59	21	38	26
	R1-12.47-2	mod. suburban & lt. rural	2.83	264	4.5	10.3	2.00	2.04	2.82	30	30	30
	R1-12.47-3	moderate urban	1.35	22	8.0	1.9	1.27	1.25	1.60	10	10	8
	R1-12.47-4	heavy suburban	5.30	50	4.0	2.3	4.31	4.09	5.65	12	17	12
	R1-25.00-1	light rural	2.10	115	6.0	52.5	2.35	2.23	3.00	28	23	30
	R3-12.47-1	heavy urban	8.40	472	12.0	4.0	6.64	6.30	8.70	20	31	25
	R3-12.47-2	moderate urban	4.30	62	14.0	5.7	3.45	3.27	4.40	13	22	18
	R3-12.47-3	heavy suburban	7.80	1,733	4.0†	10.4	7.54	7.00	9.67	56	48	55

\* Climate region of origin is indicated by R1 (temperate west coast) or R3 (arid southwest). Nominal voltage is designated by 12.47 or 25.00 (kV).

† Changed from default of 7.0 kW due to an excess of streetlighting. See [4, 67] for the relationship between avg. house size and street lighting.

In figures, shape indicates Berkeley (■), Los Angeles (●) and Sacramento (▲) results. Black symbols with dashed lines show means for each location.

**4.3.1.2.1 Feeder Topologies** Pacific Northwest National Lab (PNNL) has compiled a set of representative “taxonomy” feeders drawn from utilities throughout the United States [67]. These feeders are organized by climate region. For this work, we selected the eight feeders originating from region 1 (temperate west coast) and region 3 (desert southwest) as these climates dominate California. Table 1 presents a summary of the selected feeders.

**4.3.1.2.2 Locations and Timeframe** We simulated each of the eight feeders in three locations – Berkeley, Los Angeles and Sacramento – during the 366 days between September 25, 2011 and September 24, 2012, inclusive. We chose these locations and time span due to the availability of high-resolution PV generation and weather data. See Sections 4.3.1.2.4 to 4.3.1.2.6 for more on this data and feeder placement. Note that the California peak demand during the selected year was fairly typical relative to the past decade, with a peak load of 46 846 MW in 2012 versus a high of 50 270 MW in 2006 [10]. This means that the simulations do not include extreme conditions that may affect PV’s overall value in important ways in the long run.

**4.3.1.2.3 Feeder Loads and Power Factors** Because the taxonomy feeders specify only static planning (i.e. peak) loads, PNNL provides a script to populate the feeders with time-varying residential and commercial loads [4]. The loading process is discussed in detail in [68]; we limit the discussion here to a few points of relevance.

The PNNL method models end-use loads with “house” objects that have a weather-dependent HVAC component and schedules for other types of loads such as appliances. The schedules for each house are scaled and time-shifted to provide heterogeneity among loads. Commercial loads are modeled as groups of “houses” with a different set of load schedules corresponding to commercial activities. There are no industrial or agricultural loads.

The PNNL script applies a different distribution of load types depending on the climate region selected; e.g. air conditioning is more common in region 3 than in region 1. In this study, we applied region 3 loads to Los Angeles and Sacramento simulations and used region 1 loads in Berkeley, in keeping with the actual climate zone location of these cities.

Referring to the literature [23, 68, 71], we adjusted the script-default load power factors as summarized in

Table 2: Power Factors by Load Type

HVAC		Residential		Commercial	
Base HVAC	0.97	Water heater	1.0	Int. lights*	0.90
Fans	0.96	Pool pump*	0.87	Ext. lights*	0.95
Motor losses	0.125	Other res.*	0.95	Plug loads*	0.95
				Street lights	1.0

\* Power factor was changed from the PNNL default value of 1.0.

Table 2. We also reduced a capacitor bank on one feeder (R1-25.00-1) from 150 kvar/phase to 50 kvar/phase after noticing that it was overcompensating for reactive power, possibly because it is a rural feeder and is meant to handle more pumping load.

**4.3.1.2.4 PV Generation Data** The PV integrator SolarCity provided us with a database of instantaneous power at each inverter they monitor (roughly 7,000 systems, mostly in California) under the terms of a non-disclosure agreement. The vast majority of inverters provide data on the quarter hour; some have one-minute data for varying portions of our timeframe.

We performed data quality filtering to ensure we used only complete and credible profiles in our models. We discarded 448 individual anomalously high recordings (greater than 125% of the rated capacity of the installation). We then dropped profiles that had more than 3.3% of data missing between 8:00 and 16:00 or spurious non-zero night time readings. In Berkeley, we started with 603 profiles, of which 325 (54%) passed our data quality checks. In Los Angeles, 187 out of 355 (53%) passed, and in Sacramento, 308 out of 465 (66%).

To address remaining missing readings in the selected profiles, we selected a very complete profile (with at least 365.8 days of non-zero readings between 8:00 and 16:00) from near the center of each location. We used readings from these “filler” profiles to fill gaps longer than one hour in other profiles from that location, scaling the filler readings by the ratio of the two profiles’ rated capacity. Any shorter gaps we allowed to be handled internally by GridLAB-D, which uses the last-seen generation value until the model clock reaches the timestamp of the next reading.

**4.3.1.2.5 Weather Data** We obtained one-minute temperature, humidity, and solar irradiance data for Berkeley from Lawrence Berkeley National Laboratory [38] and for Los Angeles and Sacramento from SOLRMAP at Loyola Marymount University and Sacramento Municipal Utility District [6]. The Los Angeles and Sacramento data, having been quality controlled at the source, appeared to be quite complete and reliable and was used with only minor reformatting. The Berkeley data, consisting of “raw” weather station measurements was somewhat less pristine and required the following edits:

1. The data did not include a reliable measurement for direct solar irradiance, but we were able to calculate this from measurements of global and diffuse solar irradiance using the solar zenith angle.
2. Sometimes measures of diffuse and global irradiance were missing or zero during the daytime when true darkness was unlikely. For periods like this of one hour or less we simply interpolated between the measurements at the beginning and end of the missing data; we did this for a total of 30 hours throughout the year. For longer gaps, we copied in data from nearby days in our data set; we

selected days with similar cloud conditions according to observations from Oakland International Airport, 18 km (11 mi) south [11], to maintain reasonable consistency in overall patterns of irradiance. A total of 37.4 days were substituted in this way, with a concentration in mid-December 2011.

3. Some temperature data was missing as well. We filled 5.5 days of gaps less than one hour long using the same interpolation method used for solar irradiance. We filled 25.6 days worth of longer gaps with hourly temperature observations from Oakland International Airport [11].

The weather data determined HVAC load in GridLAB-D but was not used to simulate PV generation, which was instead extracted from the SolarCity database. By using generation data points located not far from the weather stations – at most 39 km (24 mi) distant in Berkeley, 27 km (16 mi) in Los Angeles and 45 km (28 mi) away in Sacramento – we preserved some (if not all) of the correlation between air conditioning load and PV generation. Given that buildings have significant thermal mass (resulting in a lagged and smoothed response to weather) and our goal was to preserve broad correlations between PV output and building load, we believe that the necessary corrections to the Berkeley weather data are acceptable and do not substantially affect our results.

In Berkeley, the low, mean and high temperatures for the year were 0 °C (32 °F), 13 °C (56 °F), and 35 °C (94 °F), respectively. In Los Angeles, the low, mean and high temperatures were 4 °C (39 °F), 17 °C (62 °F) and 34 °C (94 °F) and in Sacramento they were –4 °C (25 °F), 16 °C (61 °F) and 43 °C (109 °F).

**4.3.1.2.6 Geographic Assignment of PV Profiles** We sought to attach PV profiles to GridLAB-D houses in a way that reflects the diversity of solar generation over the area of a distribution feeder. This geographic diversity is driven in part by variations in cloud cover, but also by differences in PV system orientation, technology and shading – all of which are reflected in the SolarCity data set.

The GridLAB-D taxonomy feeders are anonymized and therefore we do not know their physical layout. However, the models do contain electrical connectivity for all components and lengths for each overhead and underground line segment. We used this information and the graph layout utility Graphviz to create a geographic layout for each feeder subject to these constraints. Images of these layouts are available online [30].

We then used ArcGIS to superimpose the resulting feeder layouts on the SolarCity profile sources. We manually placed the feeders in locations with high densities of generation profiles to capture as much spatial diversity as possible. We then ran a “nearest neighbor” query to assign each distribution transformer to the closest SolarCity profile with acceptable data quality. 97 profiles in Berkeley, 99 profiles in Los Angeles and 101 profiles in Sacramento were matched to at least one transformer. Table 1 lists the number of profiles used for each scenario.

**4.3.1.2.7 Penetration Levels and PV Placement** For each GridLAB-D run, we populated only a portion of the houses with PV, to test various levels of penetration. To define “penetration” we first needed to establish a baseline loading for each feeder. To this end, we executed a baseline run for each feeder (with no PV) in each location and recorded its peak load. We then defined penetration as:

$$\text{PV penetration} = \frac{\sum (\text{PV system ratings})}{\text{Peak feeder load from baseline run}}$$

We tested PV penetration levels of 0%, 7.5%, 15%, 30%, 50%, 75% and 100%. We placed PV randomly across the available house models and used the same random number seed for all scenarios. Using the same seed ensured, for example, that houses populated with PV in the 15% penetration scenario were a strict subset of those populated in the 30% penetration scenario, isolating as much as possible the effect of penetration from the effect of placement. We used the same random ordering of houses for PV placement in our three test locations. We modeled the PV as a unity power factor “negative load”.

All penetration levels should be treated as approximate for two reasons. First, our denominator for penetration was the baseline peak load during our test year, rather than the long-run feeder peak load which would typically be used in situations where more data was available. Second, due to transformer scaling (see Section 4.3.1.2.9) and other minor adjustments, the peak loads from our final 0% penetration runs differ slightly from the peak loads of our baseline runs. In general this difference is small, with the 0% penetration runs having peak load ranging between 3.9% lower and 2.9% higher than the baseline runs. However, in one scenario (R1-12.47-3, Berk.) the final peak load was 8.0% lower than the baseline peak load. So in this worst case scenario the nominal 100% penetration might more accurately be read as a 108.7% penetration.

**4.3.1.2.8 PV Generation Profile Scaling** All of our selected PV generation profiles appear to be residential-scale, with system ratings ranging from 1.68 kW to 13.16 kW. To establish a reasonable installation capacity for each building, we first used the following formula from PNNL’s load population script [4]:

$$\text{building PV rating estimate} = A \times 0.2 \times 92.902$$

where  $A$  is the floor area of the building in square feet, 0.2 is a rough estimate of the rated efficiency of the installations, and 92.902 W/ft<sup>2</sup> is the “standard test conditions” insolation.

We scaled up all commercial PV generation profiles so that their ratings matched this rating estimate. For residential installations, we scaled down the generation profile if its rating was *higher* than the rating estimate for the house. We did not scale up residential profiles with ratings smaller than the rating estimate since it is common for residential installations not to occupy the entire roof space.

**4.3.1.2.9 Transformer Scaling** Transformer aging is one of our outcomes of interest, and it depends not on absolute loading of the transformer but loading relative to the transformer’s rating [2]. While our simulated loads are roughly scaled to the planning load value listed at each transformer in the taxonomy feeders, our loads may be somewhat larger or smaller than the planning loads due, for instance, to our use of different weather data at our three locations. This means that, unless corrected, some transformers would be sized inappropriately for the loads attached to them.

To address this issue, we assembled a “menu” of distribution transformers in standard kVA sizes based on the units present in the taxonomy feeders and manufacturers’ data [1, 3]. We then replaced each transformer with the smallest transformer from the menu with a rating greater than the observed peak apparent power for that transformer from the baseline run. This is a conservative size estimate for distribution transformers given that in practice many carry power over their ratings during peak periods [2].

Note that to some extent the concern about transformer sizing also applies to conductor sizing; some taxonomy feeder line conductors may not be sized appropriately for our simulated loads. Because

conductor sizing was not a focus of this work, we did not undertake to resize the conductors in the way we did the transformers, and indeed when we run GridLAB-D we occasionally observe warnings that conductors are modestly overloaded. This may slightly distort our absolute results for line losses. To address this we instead report the percent change in losses between penetration scenarios. The percent change should not be affected significantly by conductor size since line resistance is a linear scaling factor on line losses and all penetration levels use the same conductors.

**4.3.1.2.10 Voltage Regulation Setpoints** All of the taxonomy feeders have an on-load tap changer (LTC) at the substation, and two of them feature additional line voltage regulators. During the baseline runs, we observed that the upper bound of the LTC and regulator deadbands were set at approximately 1.05 pu, right at the edge of ANSI standards for end-use voltages. This contributed to a significant number of voltage violations due to time lag in regulator response when voltages rose outside the deadband. We therefore lowered the top of the LTC and regulator deadbands to 1.04 pu (maintaining the bandwidth) for our production model runs.

**4.3.1.3 Physical Results** Our results must be interpreted with several important caveats in mind. First, our simulation covers one particular year that was chosen primarily for PV data availability. It may not include extreme weather or other events that would drive true system peaks in the long term. Second, though the GridLAB-D load models are physically-based and the taxonomy feeders are based on real feeders, we have not modeled the actual feeders and loads in our study locations. Third, the prototypical feeders are “typical”, meaning they do not have special problems such as poor voltage regulation or capacity constraints that would require special attention when integrating PV.

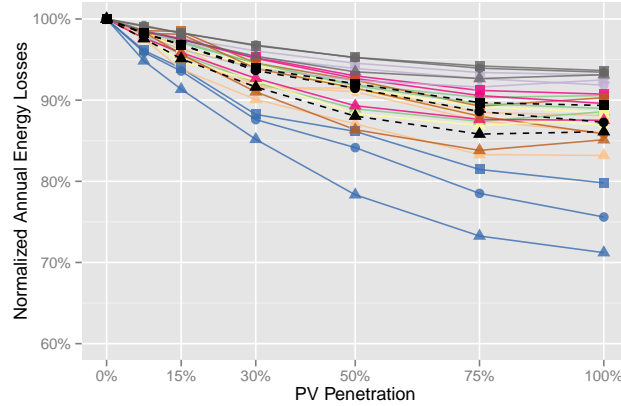
The base quantity for all normalized results is the value of the metric in question for the feeder at 0% penetration.

**4.3.1.3.1 System Losses** We measured instantaneous system losses (including transformer and line losses) every fifteen minutes. As shown in Figure 1a, we found that increasing PV penetration decreased system losses, with diminishing effects at high penetrations. The impact of PV on losses was similar across the three locations, but varied considerably by topology, with losses reduced by anywhere from 7% (R3-12.47-3) to 28% (R1-25.00-1) at 100% penetration.

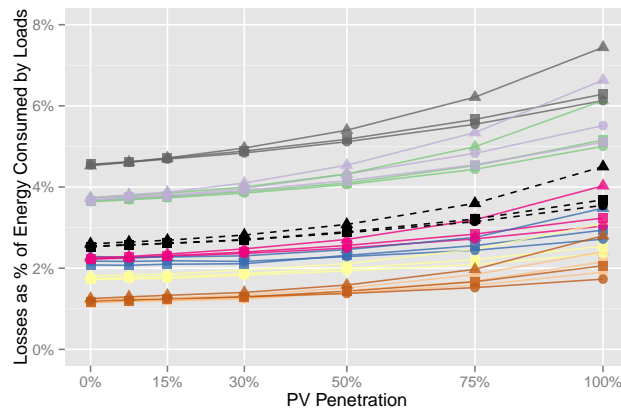
We attribute the reduced marginal effect of PV at high penetrations to the fact that losses are proportional to current squared; the more PV reduces power (and thus current) flow on the lines, the less effect further reductions will have on losses. For some feeders (mainly in Sacramento) losses *increased* as penetration rose from 75% to 100%, presumably because the losses associated with high “backflow” currents at certain times began to exceed the losses “saved” at other times when net current flow was lower.

Figure 1b shows that losses as a percentage of energy consumed by loads from the grid (i.e. as a percentage of utility wholesale power purchases) generally increase with PV penetration. This is likely because most of the net load reduction happens off-peak, when system losses are lower than on-peak.

**4.3.1.3.2 Peak Loading** We measured peak load as the maximum fifteen-minute rolling average of one-minute measurements at the substation. The extent to which PV reduces feeder peak load depends



(a) Normalized system losses.



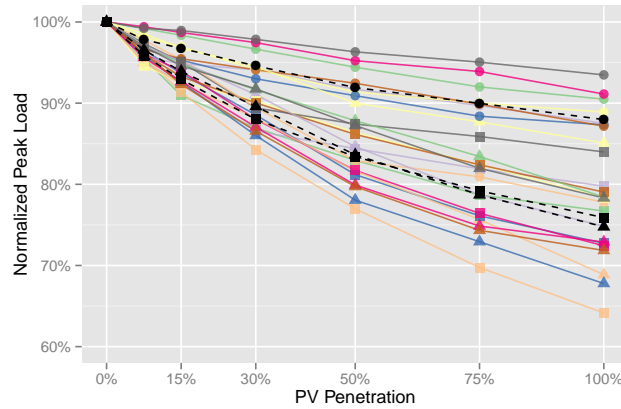
(b) Losses as a percentage of load energy supplied by the grid.

Figure 1: System losses. See Table 1 for key.

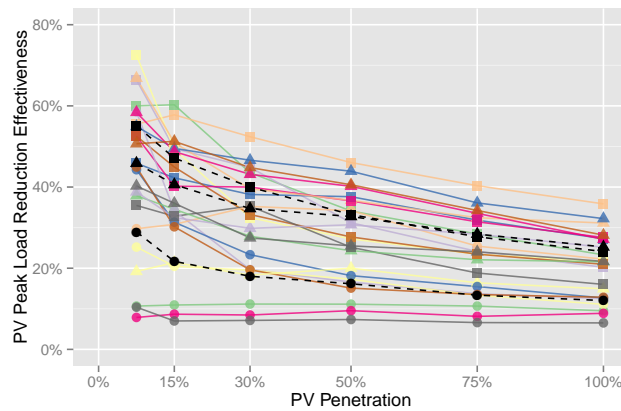
largely on the timing of the peaks. Clearly, peak load reduction will be greatest if peak load is coincident with peak PV production. In California, however, load typically peaks later in the day than PV production, and therefore peak loads are reduced by only a fraction of the PV's rating.

As shown in Figure 2, we observed that PV generally reduced peak loads by much less than the penetration percentage. In contrast to system losses, location (i.e. climate) had a strong effect on the peak load reduction impact of PV, with Sacramento and Berkeley showing more significant reductions than Los Angeles. Figure 2a shows the normalized peak load as a function of PV penetration, whereas Figure 2b shows the peak reduction as a percentage of the solar penetration. Figure 2b illustrates that low penetrations of PV can be quite effective at reducing peak loads, although this is not true in all cases. Peak load reduction effectiveness diminishes as penetration increases because early increments of PV tend to reduce daytime peaks, causing the new peak to be in the evening when PV contributes less power.

Figure 3 illustrates trends in the timing of peaks as PV penetration increased. Without PV, peak loads arrived in August 2012 for most Sacramento feeders and half of the Los Angeles feeders, while Berkeley feeders generally peaked in fall 2011 or June 2012. Peak times were widely dispersed between 14:22 and 17:18. However, a 7.5% penetration of PV was sufficient to eliminate August peaks for all but one Los



(a) Normalized peak loads.



(b) Ratio of peak load reduction to penetration

Figure 2: Effect of PV on peak loads.

Angeles feeder, shifting their peaks to the later afternoon during a relatively warm spell in October 2011. Berkeley peaks, while initially shifting towards the summer, were ultimately also moved to the fall by high penetrations of PV. Meanwhile the Sacramento peaks, driven by larger air conditioning loads, remained in the summer at all levels of penetration, although moving noticeably later in the afternoon. In all locations, peaks were moved later in the day as PV reduced daytime usage.

When interpreting our peak load reduction results, it is important to consider how well the simulated feeder load shapes align with feeder load shapes actually found in California. We do not have access to a large enough corpus of load shapes to do a rigorous analysis of this issue, but a high-level comparison will suffice to contextualize our findings. Figure 4 shows the average hourly load and PV generation for each of our simulated feeders on August 13, 2012, which was the day CAISO recorded its peak demand for 2012 [10]. It is also the peak demand day for five simulated Sacramento feeders, though not for any Los Angeles or Berkeley feeders. Each individual profile is normalized against the peak hour for that profile. As in the other figures, the locational means are straight averages of the eight normalized feeder simulations, i.e. the feeders are not weighted by their size or expected frequency of occurrence in the field. The load plot also shows normalized CAISO system load (larger green circles) and PG&E system load (larger blue circles).



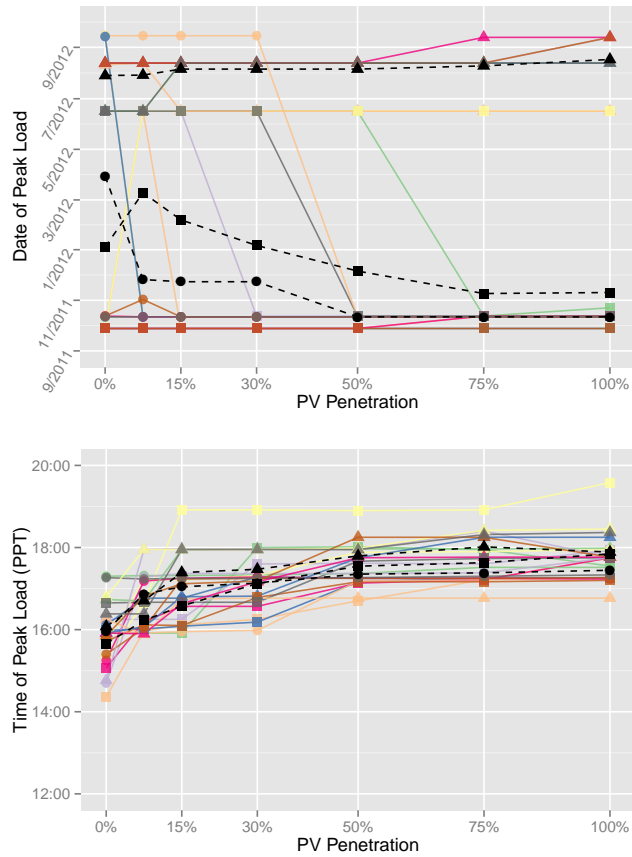


Figure 3: Date and time of peak loads. The time reported is the first minute of the peak fifteen-minute period.

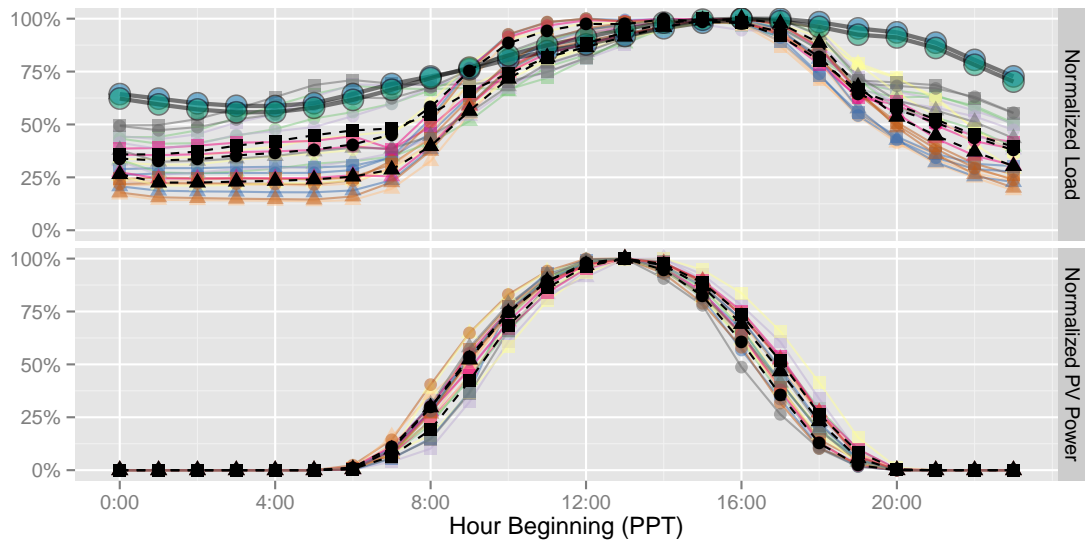


Figure 4: Normalized hourly load and PV generation profiles for August 13, 2012. Normalized PG&E system load is shown by larger blue circles and CAISO load by larger green circles [9].

From this figure we can see that the simulated peaks match well with the PG&E and CAISO peaks in the

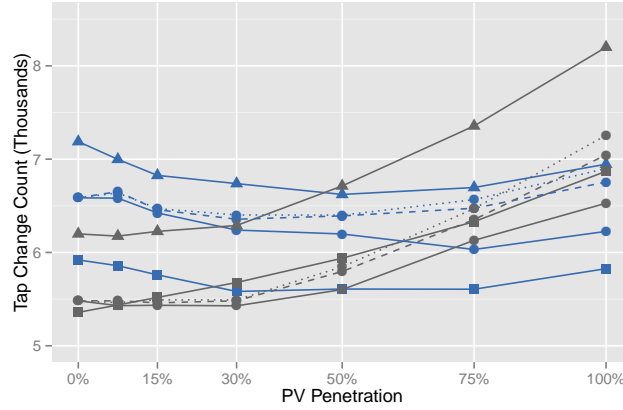
15:00-16:00 range. However, the simulated feeders universally drop in demand more quickly than the system in the evening. This has significant implications for peak load reduction. Note from the bottom panel in Figure 4 that the last hour in which PV can be expected to have a noticeable impact on load on this day is 19:00. A feeder with a very high PV penetration might therefore be expected to end up with a daily peak in the 20:00 hour. At this time, the simulated feeders are generally running at 40-75% of their daily peak load, whereas the CAISO system is still at 91% of peak.

This simple one-day example ignores several factors that are important when calculating annual peak demand reduction, such as load variation within each hour and the fact that PV often shifts the peak to a different day, rather than a different time on the same day. Also, the comparison to an overall system load profile greatly obscures the wide variation of individual feeder profiles that comprise it. For instance, SCADA data provided by PG&E under the terms of a nondisclosure agreement indicates that on August 13, 2012 the most common hours for feeders to peak were 16:00 and 17:00, but each of these hours only accounted for about 16% of feeders, with 37% peaking earlier (10% before noon) and 31% later in the evening [29]. Thus, it is likely that the simulated load shapes are a good match to some subset of California feeders and therefore the reported peak load reduction is achievable in some locations. However, the fact that the simulated feeder profiles are not a good match for the general system profile in the evening indicates that it would be optimistic to expect the simulated peak load reduction to occur universally across California.

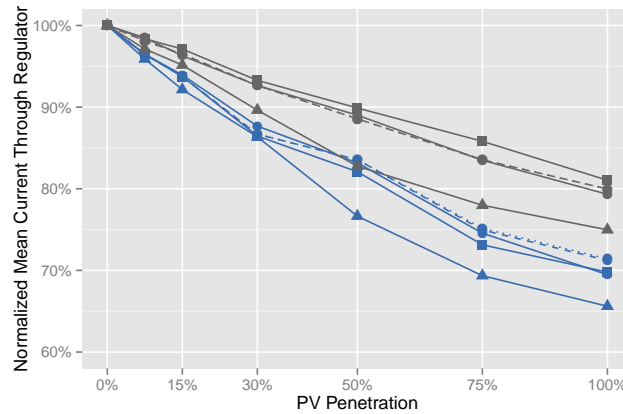
**4.3.1.3.3 Transformer Aging** GridLAB-D 2.3 implements the IEEE Standard C57.91 Annex G [2] method for estimating transformer insulation aging under various loading conditions. GridLAB-D implements the method for single phase center tapped transformers only. This is the most common type of transformer on the taxonomy feeders, but one feeder (R3-12.47-2) did not have any so it was excluded from the aging analysis. In the Annex G model, a “normal” year of aging corresponds to the amount of insulation degradation expected if the transformer hot spot were at a constant 110 °C throughout the year. A transformer that is often overloaded will age more than 1 y in a year, and one that is loaded below its rating will age less.

In general, we observed minimal aging in all scenarios and penetration levels, with a mean equivalent aging of up to 0.29 y in one scenario (R3-12.47-3, Sac.) and all other scenarios having mean aging less than 0.001 y. We attribute this slow aging to the fact that the transformers were conservatively sized at or above their baseline peak load (see Section 4.3.1.2.9). However, in R3-12.47-3 (Sac.) at PV penetrations of 30% and above we did observe a small number of transformers aging quite rapidly, up to 166 y during the simulated year. These transformers are likely at a location where net PV generation is often higher than the load they were sized to handle, and in reality they would need to be upgraded to handle this backflow.

**4.3.1.3.4 Voltage Regulators** Tap-changing voltage regulator wear and tear is driven primarily by the number of tap changes the device must perform and the current that it handles during operation. In our simulations, tap changes at the substation LTC were not affected by topology, climate or PV penetration, varying by only 1% across all model runs. This is because the models did not include a transmission impedance component, with the transmission voltage instead following a fixed schedule of values recorded from an actual substation in WECC. The substation LTC operates to maintain voltage immediately downstream within the deadband despite fluctuations in the WECC schedule, and is insensitive to downstream changes in load. Due to the lack of a transmission model, our simulations do not provide reliable insight on LTC response to PV.



(a) Tap changes.

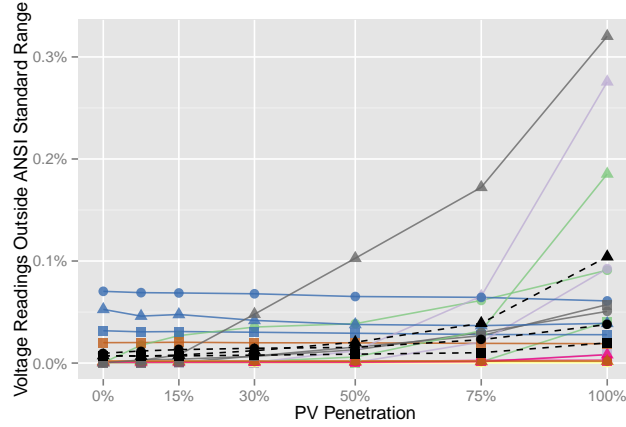


(b) Mean current duty.

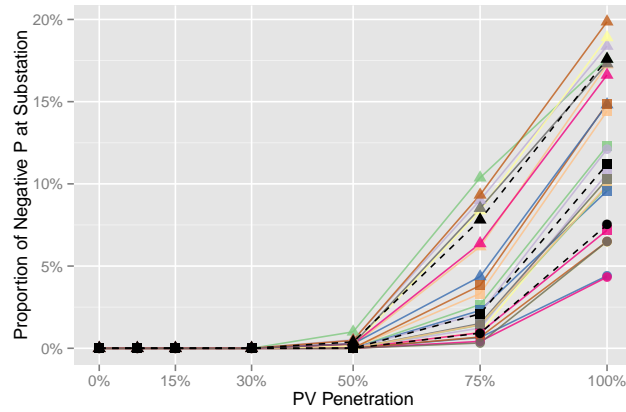
Figure 5: Line voltage regulator activity across all three phases. See Section 4.3.1.3.4 for discussion of broken lines.

The two mid-feeder regulators in the simulation (at R1-25.00-1 and R3-12.47-3) do have simulated impedances and varying loads both upstream and downstream and thus exhibit more variation. Figure 5a shows that PV has little effect at R3-12.47-3 until 50% penetration, at which point tap changes begin rising noticeably. This result echoes other work [57] and concerns from utilities that PV variability will increase regulator maintenance needs. However, the number of tap changes tends to decrease at R1-25.00-1, at least to a penetration of 50%. This decrease could be due a reduction in diurnal range of net load, and therefore the range of voltage drops. Though further investigation is needed to fully understand these results, they do indicate that in some cases PV could reduce voltage regulator maintenance needs.

We examined two sensitivity scenarios to better understand the impact that our PV data had on our regulator results. The dotted lines in Figure 5 show regulator response when we used the single PV profile with the most one-minute data available (82% of days) at all PV sites rather than using the usual geographic assignment. The dashed line shows the same scenario with the one-minute data downsampled to fifteen-minute resolution; this intermediate scenario helps us to distinguish the effect of the one-minute data from the effect of eliminating geographic diversity. We limited the sensitivities to Los Angeles because this was our source of one-minute data. Figure 5a suggests that geographic diversity reduces tap change frequency (because the solid lines which include geographic diversity fall well below



(a) Proportion of voltages outside ANSI standards.



(b) Incidence of negative real power flow through substation.

Figure 6: Voltage control and backflow. Many scenarios overlap near 0.0%.

their corresponding single-profile dotted and dashed lines) and that fifteen-minute PV data is a reasonable proxy for one-minute data when studying regulator behavior (because the dashed lines track their corresponding dotted lines closely). It is possible that with PV data on even finer time scales (less than one minute) a different pattern of regulator activity would emerge. However, since regulators generally have a response lag on the order of 30 s, very brief fluctuations in PV are likely to result in voltage changes on the feeder rather than increased regulator activity.

The effect of PV on regulator current duty was more consistent than the effect on tap changes, as illustrated by Figure 5b. With PV reducing the downstream load, current through the regulator declines steadily as penetration increases. This suggests that even in cases where PV increases a regulator's activity, its expected lifetime may stay the same or even increase because each tap change is less destructive under lighter current duty. Our sensitivity runs suggest that neither geographic diversity nor the use of one-minute resolution data has a substantial effect on regulator current duty.

**4.3.1.3.5 Voltage Quality** We recorded voltage at all point-of-use meters at fifteen minute intervals and tabulated in Figure 6a the proportion of readings falling outside of the ANSI standard range of 0.95 pu–1.05 pu. In general, voltages appear to be well-controlled, with most runs having less than 0.002% of readings out of range, and the worst case (R3-12.47-3, Sac.) having 0.32% of readings out of range. This is consistent with prior work suggesting that many feeders can support high penetrations of PV without voltage violations [45]. Except at feeder R1-25.00-1, almost all out-of-range voltages observed were greater than 1.05 pu. As expected these high-side excursions generally become more frequent as penetration increased and the power injection from PV raised some voltages locally. At R1-25.00-1 the out of range voltages were predominantly less than 0.95 pu, with a small amount greater than 1.05 pu. Under these conditions, increasing PV penetration improved voltage quality on the feeder by boosting some local voltages that would otherwise be low. As noted in Section 4.3.1.3.4, it is possible that more brief voltage excursions would be observed with higher resolution PV generation data.

**4.3.1.3.6 Reverse Power Flow** Figure 6b shows the incidence of negative real power flow (“backflow”) through the substation, which can be a proxy for protection issues and higher interconnection costs. At 50% penetration, twelve of the 24 scenarios exhibited occasional backflow, up to 1% of the time each. At 100% penetration, all scenarios experienced backflow at least 4% of the time. In general, backflow was more prevalent in Sacramento because PV penetration in Sacramento was measured against a higher peak air conditioning load. This led to a larger absolute quantity of PV generation in Sacramento but with similar low loads to Los Angeles and Berkeley on cooler days.

**4.3.1.3.7 Preliminary Observations Regarding Electrical Storage** After concluding our modeling of feeders with PV, we turned our attention to electric storage to see how it might enhance the benefits or mitigate the drawbacks of PV on the distribution system. We implemented a customer peak shaving storage control algorithm – currently being tested by SolarCity in the field – in GridLAB-D’s inverter module. We tested a variety of penetrations of storage in combination with PV, with the storage deployed at the same buildings as the PV.

Our preliminary results suggest that a customer peak shaving algorithm can be helpful in reducing feeder peak loads, with its effectiveness varying depending on the timing of individual customer loads and PV output relative to the feeder peak. Storage also somewhat reduces backflow in high PV scenarios by using some excess PV power for charging, while it slightly increases the annual energy consumed by the feeder due to inverter and battery losses during charging and discharging.

In the course of analyzing our storage results in more detail, we identified some problems in the model setup that discourage us from drawing any more specific conclusions or including storage in our economic analysis. The main issue is related to the sizing of the battery systems. We initially sized the batteries to be equal in power capacity to the PV system at the site, on the grounds that it would be economical for the battery and PV to share the same power electronics. However, given that our sizing cap for PV systems is fairly generous, especially at commercial sites (see Section 4.3.1.2.8), this resulted in batteries many times larger than would be needed for a typical peak shaving application. In extreme cases, this inappropriate sizing triggered some aberrant behavior from the peak shaving algorithm that would not occur under the conditions it was designed for. In less extreme cases, the effect was simply that the battery was able to keep the metered load almost completely flat by charging from PV during the day then discharging during the evening and at night, which is quite different from the usual understanding of “peak shaving”.

Although many of the batteries were sized appropriately and exhibited more traditional peak shaving

behavior, we believe that on the whole too many behaved inappropriately for our simulations to be a good model for widespread peak shaving storage deployment. We hope to return to this issue in future work not only by sizing the batteries more appropriately but also by analyzing and experimenting with the SolarCity algorithm to understand how its parameters can most effectively be tuned for various customers.

**4.3.1.4 Economic Results** In this section we present estimates of the economic effect that PV could have on California's power grid. Our analysis focuses on PG&E since PG&E shared with us useful data that allowed us to make better-supported estimates than we would be able to for the remainder of California. For reference, PG&E serves approximately 37% of California load, so our findings for PG&E can be scaled up accordingly to provide rough estimates for all of California. Of course, scaling in this way assumes that conditions in the rest of California are similar to conditions in PG&E territory, which may not be the case.

**4.3.1.4.1 Distribution Capacity Value of PV** Increased distribution-level peak load creates a need for investment in substation equipment such as larger transformers. Since these investments are mainly driven by the need to accommodate the peak load anticipated on the feeder, PV can defer and/or abate some of them by reducing peak load, thus creating a financial benefit. In this section we present a model that combines our peak load reduction simulation results (Section 4.3.1.3.2) with distribution system data provided by PG&E to estimate the system-wide capacity benefit of PV.

**Projects and Feeders** Our general approach to modeling changes in PG&E's distribution capacity expenses involved establishing a baseline of the number of distribution capacity projects likely to occur in the next ten years and assuming that these projects accounted for most of PG&E's distribution capacity budget (see Section 4.3.1.4.1). We then calculated the deferral of individual projects under various PV penetrations based on our normalized peak load reduction results from Figure 2a and considered the difference in the total net present value (NPV) of projects with PV to be the capacity benefit of PV at that penetration.

We assumed that a project is scheduled for a feeder in the year when the feeder's peak load reaches 100% of its rated capacity. In reality some projects are scheduled before peak load reaches feeder capacity, while in other cases a project may be deferred even though a small overload is forecast in order to better assess the true load growth trajectory of the feeder. Nonetheless, we make the assumption that in expectation projects will be scheduled in the first year that load equals or exceeds feeder capacity. In practice, other engineering and logistical considerations may also influence the timing of capacity projects; for example, a project may be initiated sooner than necessary to economize on personnel and equipment in the area when non-capacity-related work is taking place. These factors are outside the scope of this project and generally independent of PV penetration.

To assess when capacity projects would occur we used feeder-level loading data from 2012 and projected load growth data for 2013-2017 provided under the terms of a non-disclosure agreement by PG&E [29]. For years beyond 2017 we carried the 2017 growth rates forward, which is clearly a very rough prediction of future trends. Also note that in practice, PG&E relies on growth forecasts based on one-year-in-ten weather data (that is, based on the assumption of an unusually hot year) to be conservative in their capacity planning. We used projections based on one-year-in-two weather data for two reasons. First, the one-year-in-ten projections created a discontinuity between 2012 (actual loading) and 2013 (projected

loading). Second, for tractability our model is based on PV's effect on "actual" peak loading (under our modeling assumptions) rather than the full, more complex project forecasting and planning lifecycle.

In selecting PG&E feeders to run through the model, we narrowed the data set by eliminating feeders operating at 4 kV, those already having greater than 10% PV penetration, and those that had already reached 100% of their rated capacity in 2012. After eliminating these, we estimated which feeders would require a capacity project in the next ten years. This left us with about 300 feeders (roughly 10% of the 3000 feeders in PG&E network), or about 30 distribution projects per year. PG&E confirmed that this is a reasonable number of feeders to reach capacity each year, although in reality some of these situations are addressed without requiring a full-fledged capacity project, e.g. by switching load from one feeder to another, which was a nuance we were unable to account for in the model.

Note that this model considers only projects triggered by individual feeder (or "circuit") capacity constraints. In practice, PG&E also initiates projects on the basis of substation transformer (or "bank") capacity constraints (where an individual bank may serve multiple feeders) but we were unable to include this in the model because the GridLAB-D taxonomy feeders are modeled individually and we do not have any information about other feeders on their banks. Because we scale our final results to the overall PG&E distribution capacity budget this omission is unlikely to greatly influence the magnitude of the results, but it may introduce some error if the distribution of bank loading and/or load growth is significantly different from the loading distributions for feeders.

**Financial Assumptions and Time Horizons** For our calculations of total value we look out to a ten year time horizon and find net present values in 2012 dollars using PG&E's weighted average cost of capital (WACC) of 7.6% and a combined inflation plus project escalation rate of 2.5% [29]. The model looks out over a number of years to account for the inherent "chunkyness" of capacity investments – that is, looking at any particular year could give misleading results due to a lack of knowledge about how long an "avoided" project was truly deferred. We chose a horizon of ten years as a compromise between the desire to have the longest window possible for completeness versus the increasing unreliability of feeder load growth projections and other assumptions in the far future.

Although we limited the pool of initial projects to a ten year horizon, we continued to account for the cost of deferred projects out to 25 years to avoid ignoring the cost of projects that were simply deferred slightly beyond the ten year horizon. Any project deferred beyond 25 years we considered to be completely abated, with a cost of zero. Under our discount rate assumptions a project deferred to year 25 would have 71% of its present value discounted away in any case.

**PV Deployment Timeline** Since we modeled the effect of PV over time, it was necessary to make an assumption about the rate at which PV would be added over our ten year horizon while building up to the target penetration for each model run (7.5%, 15%, etc.). We considered several possibilities, including reaching the target penetration immediately in year one, linearly ramping up to the target penetration in year ten, and following an exponential trajectory to the target penetration in year ten. The first two possibilities we discarded as being too fast relative to real-world deployment trends. The exponential possibility at first seemed to be an attractive analog to the accelerating deployment of PV in California, but we encountered some difficulty in that the shape of the exponential ramp was highly sensitive to the initial condition in 2012, and furthermore our nominal initial condition was 0% penetration, which is not

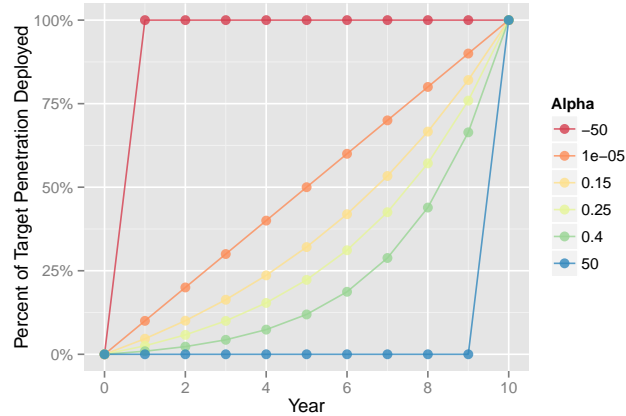


Figure 7: Representative realizations of our deployment ramp up function  $p(t)$  for varying  $\alpha$ .

possible for an exponential function. Thus, we settled on a compromise of:

$$p(t) = \frac{e^{\alpha t} - 1}{e^{\alpha T} - 1} P$$

where  $p(t)$  is the penetration in year  $t$ ,  $P$  is the target penetration,  $T$  is the year in which to reach the target penetration (ten, in our case) and  $\alpha$  is a shape parameter.  $\alpha$  values close to zero give a nearly linear ramp up, large positive values give a late ramp up, and large negative values result in a rapid rise in the early years followed by a plateau. We believe that  $\alpha$  values in the general range of 0.15-0.4 are most reasonable, with slow growth in the early years and faster growth near the end of the modeled period, but we present results for several  $\alpha$  values for comparison. Figure 7 illustrates how the shape of  $p(t)$  changes with varying  $\alpha$ .

Naturally, most values of  $p(t)$  did not correspond exactly to penetration levels that we had modeled; e.g. on the way to 15% penetration in year ten the function passes through 0.7% in year one, 1.5% in year two, and so on. In these cases, we obtained an intermediate peak load reduction by interpolating linearly between the peak load reduction of the two nearest penetrations that we had modeled.

**Applying Model Runs to PG&E Feeders** Because different simulated scenarios (feeder/climate combinations) responded differently to PV, we needed to devise a reasonable method to apply the results to the PG&E feeder data. To begin with, we limited the simulation results used to temperate west coast (R1) feeders modeled with Berkeley weather and desert southwest (R3) feeders modeled with Sacramento weather on the grounds that these were most representative of the climates the feeders originated from and best matched conditions actually found in PG&E territory. We permuted each R1 result to each feeder in PG&E’s “coastal” service territory and each R3 result to each feeder in PG&E’s “interior” service territory. (The coastal and interior regions together comprise PG&E’s entire territory.) Within each region we did not attempt to match the characteristics of the simulated feeders (e.g. peak load, length) to individual PG&E feeders, but rather assumed that any simulation result could be matched with all feeders within each region.

After calculating PV’s effect on project timing for each individual taxonomy/PG&E feeder combination, we weighted results within each territory by the regional taxonomy feeder frequencies from [67], to account for the fact that some taxonomy feeders represent more of the feeder population than others. In addition to the feeder types we modeled, The PNNL regional breakdown contained General Industrial Case (GC)



feeders – found in any region – that were not modeled in GridLAB-D and represented 10% and 20% of the R1 and R3 feeder populations, respectively. We elected not to model these feeders because they consist essentially of one large industrial or commercial load, and therefore the effect of PV on the feeder is highly dependent on the load shape chosen for this load. Without a reliable library of industrial load shapes we did not feel that we could obtain meaningful results for this feeder topology.

In lieu of modeling the GC feeders, we incorporated two alternative GC assumptions for the remainder of the economic model. The first assumed that the GC feeder portion provided no solar benefit (“No Solar - GC”) and second assumed that the GC feeder’s benefit was equal to the weighted average from the rest of region (“Solar - GC”). We believe that Solar - GC is the more realistic of the two assumptions – indeed, to the extent that industrial and commercial loads tend to concentrate their energy use during the business day they are likely to see *greater* peak load reduction from PV than residential or mixed-use feeders – but for completeness we show results with both assumptions. Note also that the PNNL feeder taxonomy does not include any networked urban cores, which represent 5-10% of the distribution system [67], so these are implicitly assumed to respond to PV like the weighted average of their region. After computing a weighted average by taxonomy feeder for each region, we averaged the regional results weighted according to the 2012 peak load of the coastal and interior PG&E feeders, which resulted in a weighting of 36.3% coastal and 63.7% interior.

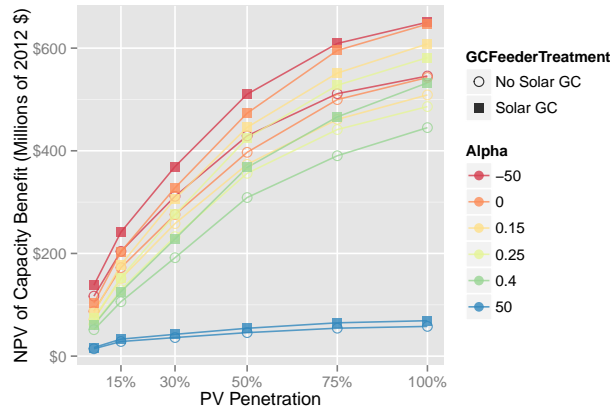
**Scaling to PG&E’s Distribution Capacity Budget** We calculated the system-wide financial benefit of project deferral by multiplying the normalized savings across PG&E’s service area with the fraction of PG&E’s distribution budget that could reasonably be affected by PV. This fraction includes the following major work categories (MWC) [5, Workpaper Table 12-5].

- MWC 06A - Feeder Projs Assoc with Substation Work
- MWC 06D - Circuits Reinforcement (DE Managed)
- MWC 06E - Circuits Reinforcement (PS Managed)
- MWC 46A - [all projects]

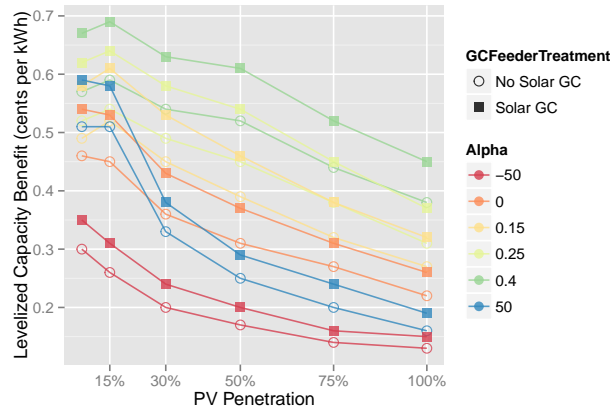
We excluded some smaller distribution expenses that would not likely be influenced by PV’s peak load reduction, such as MWC 06G, Voltage Complaints [29]. In total, 92% of PG&E’s distribution capacity budget was considered influenceable by the peak load reducing effects of PV. For 2013-2016 we used nominal budget projections directly from [5, Workpaper Table 12-5] and for 2017-2022 we used the average nominal budget for 2013-2016 before discounting back to real 2012 dollars using our standard discounting assumptions. For scale, the calculated total PV-sensitive budget in 2012 was \$132,985,000.

**Capacity Value Results** Figure 8a displays possible values of the capacity benefit across PG&E service territory given various input assumptions. Figure 8b levelizes this benefit across the kWh of PV generated throughout the ten year horizon. As expected, the capacity benefit rises with PV penetration but with diminishing returns as PV becomes less effective at peak reduction due to peaks being pushed into the evening (see Section 4.3.1.3.2).

Note that earlier deployment (that is, smaller  $\alpha$ ) always improves the NPV of the capacity benefit, but has a less consistent effect on the levelized benefit. This is because while earlier deployment defers more



(a) PG&E system-wide capacity benefit.



(b) Levelized capacity benefit.

Figure 8: Calculated capacity benefit of PV under varying assumptions.

capacity projects, it also increases the kWh generated over the ten years. If the PV arrives “too early” it may generate quite a bit of energy before its capacity actually makes a difference in deferring a project. Thus, the levelized benefit is relatively low in scenarios where PV deployment is rapid (because it provides early generation that is not yet needed to defer most projects) and when it arrives late (because it arrives after many projects have already occurred due to load growth). The levelized benefit is greatest with intermediate  $\alpha$  values, which cause the solar deployment trajectory to more closely follow the feeder load growth trajectories, avoiding projects but not generating much more energy than is necessary to do so.

**Caveats** This analysis assumes that as the peak load decreases from increasing levels of distributed solar, utilities will require less capacity equipment to service the load. From a utility perspective, uncertainty in the reliability of distributed solar in blackout or other emergency conditions may prevent some or all of the capacity benefit from being realized during the investment planning process. Aggressive capture of the capacity benefit may result in “less slack” and diversity in distribution system, which utilities may rely on in emergency or repair situations. The degree to which these phenomena affect utility planning decisions was outside the scope of this study and is an opportunity for future research.

It is also important to note from the planning perspective that a feeder projected to reach 100% of its rated capacity may result in a variety of responses with varying costs. However, normalizing the model's results and applying them to the entire distribution budget implicitly factors in the variety of responses. That is, included in the distribution budget may be some potential projects that were solved through less expensive means than a full upgrade, such as switching loads to different feeders. Our analysis implicitly assumes that the *number* of situations requiring action may change with increasing PV penetration, but the relative mix of actions taken will not change.

Recall also that all results are based only one year of simulation. Finally, the caveats from Section 4.3.1.3.2 regarding the potential overstatement of peak load reduction due to questionable load shapes mean that the capacity values reported here are likely optimistic.

**4.3.1.4.2 Energy Cost Reduction Value of PV** PV provides an energy benefit to the grid by displacing generation that would otherwise need to be procured from thermal power plants. PV can be expected to provide a value somewhat larger than the average wholesale energy price because 1) PV tends to generate during the day when prices are higher than at night, and 2) PV has a tendency to reduce distribution system losses, thus reducing overall demand by more than the energy actually generated. (PV may also reduce transmission losses, further enhancing its value, but we neglect this effect since our simulations did not include a transmission model that would allow us to gauge this change.) Note that there may also be a slight countervailing tendency for PV to *increase* demand by raising voltage and therefore causing voltage-dependent loads to consume more power.

To capture all of these nuances, we calculate the net energy benefit for a given feeder as the cost to supply the energy demanded at the substation at 0% PV penetration minus the cost to serve the substation at the given PV penetration. In other words, we calculate how much PV reduces energy costs at the substation. We obtained hourly locational marginal prices (LMP) from CAISO's day-ahead market for nodes CLARMNT\_1\_N001 (Berkeley) and WSCRMNO\_1\_N004 (Sacramento) [9]. We compared several nodes in the general area of Berkeley and Sacramento and chose these two arbitrarily after confirming that differences in price relative to neighboring nodes were very small.

Having determined the energy benefit at each feeder, we calculated a weighted average benefit within each region and then across regions using the same methodology as described in Section 4.3.1.4.1. This method provided the energy benefit for a "representative sample" of feeders across all of our simulated penetration levels. We then used an end-use consumption metric to scale the representative energy benefit to the system level. For each feeder, we calculated end-use consumption by subtracting system losses from substation energy at 0% PV penetration. We then used the same weighted average to arrive at the end-use consumption for our sample of feeders.

To find the total consumption to scale to we obtained 2012-2022 California Energy Commission (CEC) consumption growth projections for the PG&E service area. The CEC's "CED 2011 Revised-Mid" forecast predicts an average energy consumption growth rate of 1.25% between 2011 and 2022 [7, p. 6], but this figure was reduced by anticipated future PV generation and thus was not directly comparable to the end-use consumption from GridLAB-D [8, pp. 36-40]. To render the metrics equivalent, we added expected PV generation back into the 2012-2022 CEC consumption figures to arrive at overall end-use consumption. In 2012, solar generation accounted for about 1% of PG&E's consumption, as calculated from the PG&E feeder-level installed PV capacity data (assuming an 18% capacity factor, based on our simulation results) and the CEC 2011 Revised-Mid consumption forecast, projected to 2012 [7]. In 2022 we found a forecast solar generation / consumption ratio of roughly 2% using a projected 1268 MW of

installed capacity in PG&E (midway between high and low estimates in [7, p. 28]), forecast consumption of 123 353 GW h [7, p. 6] and an 18% capacity factor.

To convert the CEC consumption figures to end-use consumption, we multiplied the CEC data by one plus the solar generation ratio, scaled linearly from 1-2% over the 10 year period. Next, we determined the multiplier necessary on our sample to bring its energy consumption in line with the system's overall energy consumption. The calculated multiplier ranged from 5720 (year one) to 6453 (year ten). Using this scale factor, we were able to scale up the representative feeders' energy benefit to the PG&E system for each year out to the ten year horizon. Taking the NPV of the ten year period yields a total energy benefit comparable to the capacity benefit calculated in Section 4.3.1.4.1.

Note that the calculated multiplier was on the order of 6,000 whereas in reality there are approximately 3,000 feeders in PG&E's system. This implies that our weighted sample is unrepresentative of PG&E's system in at least one respect: the average PG&E feeder uses about twice as much energy annually as our weighted average feeder. Since the sample is being scaled to the full system size this discrepancy does not affect the overall magnitude of the results, but it indicates that the feeders (and/or weighting of feeders) chosen are not fully representative of PG&E's system, which may introduce some error.

The levelized benefit was calculated by dividing the total benefit by the NPV of PV generation for the period. We recorded PV generation from the feeder simulations and weighted and scaled them to the PG&E system using the same weights and multipliers as for energy consumption. Further note that each year's PV generation figures were linearly interpolated between simulated scenarios as described in Section 4.3.1.4.1.

At all penetration levels, we found a levelized energy benefit very near \$0.035/kWh (with a range of \$0.0349/kWh–\$0.0351/kWh). Depending on penetration level and input assumptions, this is about an order of magnitude greater than the levelized capacity benefit reported in Section 4.3.1.4.1. Because the energy cost benefit is levelized by the discounted PV generation over the ten years, it is insensitive to changes in target penetration, ramp-up trajectory (Section 4.3.1.4.1) or the treatment of the GC feeder (Section 4.3.1.4.1). The minor differences found from scenario to scenario are a result of random variations in which PV generation profiles were chosen and where they were placed with each incremental increase in penetration (see Section 4.3.1.2.7). These minor variations had a small effect on the timing of generation and the extent to which PV production reduced losses.

Note, however, that our results assume that LMPs will remain consistent for the next ten years regardless of PV penetration, which is unlikely. For reference, the average LMP during our test year was \$0.0286/kWh in Berkeley and \$0.0304/kWh in Sacramento, for a weighted average of \$0.0297/kWh. Thus, thanks to timing and loss-reduction advantages, each kWh of PV was about 18% more effective at reducing system costs than a generic resource that generates around the clock and does not reduce losses. The size of the timing premium depends greatly on the range of diurnal price variation, and it may be considerably larger or smaller in future years depending on whether California has insufficient or excess generation capacity during the day. Higher PV penetrations will tend to depress daytime demand and therefore prices, which will make further marginal additions of PV less beneficial than this analysis would suggest.

**4.3.1.4.3 Voltage Regulators and Voltage Quality** Our physical results for voltage regulators (Section 4.3.1.3.4) are sparse in the sense that 1) we do not have meaningful results for substation LTCs, and 2) only two feeders had mid-feeder regulators; these results are meaningful but the sample size is small. However, we can make some very general estimates as to how regulator maintenance expenses

might change if the trends observed in our small sample are scaled up to the system.

There are a handful of PG&E major work categories (MWC) related to voltage regulators. MWC BK (Distribution Line Equipment Overhauls) is a category that includes needed overhauls for line reclosers and regulators; in 2012 expenses of \$2,645,000 were forecast for this purpose [5, p. 5-34]. Regulators constitute about 41% of the total units of line equipment (regulators + reclosers) [29]. We assume that the unit cost to overhaul a regulator is about the same as the unit cost for a recloser and arrive at regulator overhaul expenses of roughly \$1,085,000. MWC 48 (Replace Substation Equipment) includes several “Subprograms < \$1M”, including a line item for regulator replacements projected to be \$297,000 in 2012 [5, Workpaper Table 13-16]. Some LTC replacement work also takes place under MWC 54 (Distribution Transformer Replacements) which had an overall forecasted value of \$61,005,000 in 2012 [5, p. 13-14]. However, most of this expense is for general substation transformers not LTCs, and projects are usually triggered by factors unrelated to the LTC such as dissolved gas analysis of the transformer oil; in these cases the LTC is replaced in the course of a larger project rather than due to wear on the LTC itself [29]. Therefore we conclude that MWC 54 expenses are unlikely to be affected by changes in LTC operation triggered by PV. This leaves us with a total projected 2012 regulator budget of \$1,382,000 from MWC BK and 48 that could be affected by changes in tap-change activity.

If we assume that substation LTCs will respond similarly to line regulators, we can extrapolate our regulator results to the system and estimate how much PV might affect overall regulator expenses. At the high end (R3-12.47-3, Sac.), PV increased regulator operations by 32% at 100% penetration. If this were representative of the system, maintenance expenses would also increase by 32%, or roughly \$442,000 in 2012. In a more optimistic scenario where regulator operations decreased by 8% due to the presence of PV (as at R1-25.00-1, L.A. at 75% penetration) across the system, regulator maintenance expenses might decrease by \$111,000. In reality both of these scenarios might exist somewhere in the system, in addition to many intermediate cases and a few more extreme ones, likely resulting in an overall expense change somewhere between these bookend values. The overall impact will be more favorable if the reduced current duty brought about by PV also extends regulator lifetime, but the sensitivity of regulator lifetime to reductions in current is heavily dependent on the regulator model and its pre-PV current duty, so we lack the data to estimate the magnitude of this effect. In any case, the clear conclusion of the budgetary analysis is that any regulator maintenance cost changes will be very small in comparison to the energy cost and capacity cost effects of PV.

For comparison, PG&E’s budget for addressing Voltage Complaint Projects Involving Secondary Distribution (MWC 06G) was forecast to be \$2,800,000 in 2012; some fraction of MWC 06E (Circuits Reinforcement – Project Services Managed, forecast at \$36,941,000 in 2012) is also dedicated to “primary distribution voltage correction work” [5, p. 12-20]. As noted in Section 4.3.1.3.5 voltage quality on our simulated feeders was only mildly affected by PV, although we expect that in the field there will be some marginal feeders where it will be a significant issue. Our data, however, are not sufficient to make a good estimate of how frequently PV will actually trigger complaints or create serious enough problems to require additional work in the abovementioned MWCs.

**4.3.1.4.4 Transformer Aging and Backflow/Protection** As noted in Section 4.3.1.3.3 we observed minimal transformer aging across all of our simulated scenarios, with little change due to PV except with one particular feeder/climate combination. We attribute the lack of aging mainly to the conservative sizing of the distribution transformers relative to the loads served. We attempted to locate a data set of distribution transformer loading to ascertain how well this assumption matched California’s actual distribution transformers, but it seems that utilities generally do not track the loading of these

transformers closely. This is understandable given that distribution transformers are numerous, highly dispersed, and relatively cheap to replace.

We do expect that PV will have some effect on transformer lifetimes in areas where transformers are loaded at or above capacity. In most cases, lifetime is likely to be extended as daytime transformer loading is reduced by generation on the secondary side. In some cases where the installed PV power is much greater than the previous daytime load, transformer lifetime may be decreased by large reverse power flows. Given the uncertainty about existing transformer load shapes and ages it is difficult to estimate the size of the benefit (or cost) that PV could provide.

Similarly, we refrain from drawing any conclusions about the economic effect of backflow caused by high PV penetrations (see Section 4.3.1.3.6 for physical results). The main concern regarding backflow is that it may require modifications to protection systems that were designed with only one-way power flow in mind. Determining whether such corrections are necessary on any given feeder requires a specialized protection analysis which is beyond the scope of this study.

**4.3.1.5 Conclusion** In general, our physical results suggest that PV penetrations up to 50% reduce system losses and feeder peak loads while having positive or negligible effects on transformer aging, regulator wear, and voltage quality, though there is a small incidence of backflow at 50% penetration. At higher penetrations we observed diminishing benefits for system losses and, in some scenarios, undesirable impacts on other metrics. Most effects displayed significant variability related to feeder topology and climate; in particular, both positive and negative impacts of PV at a given penetration tend to be strongest in Sacramento and weakest in Los Angeles, with Berkeley intermediate.

Our economic results indicate that the desirable timing and loss reduction associated with PV power make it about 18% more valuable than an average LMP might suggest. However, at high penetrations PV will likely itself affect prices, which will change the calculation. We also find that PV is likely to defer or avoid some otherwise necessary distribution capacity projects in California, and thus it provides some distribution capacity value, although this value is small in comparison to its energy value. We investigate some ancillary economic effects of PV and find some (e.g. changes in voltage regulator maintenance expenses) to be very small in the context of California's overall energy system, while others are beyond our ability to estimate at this time.

These conclusions must be interpreted cautiously in light of limitations in the models, such as the one year time span, the reliance on simulated loads that do not match well to California's aggregate load profile at key times of day, and the limited time and spatial resolution of the PV generation data. Further caveats are noted in the text where appropriate.

**Acknowledgment** We wish to thank John Carruthers, Donovan Currey & Matt Heling of PG&E; Jason Fuller and the GridLAB-D engineering team at PNNL; Eric Carlson, Justin Chebahtah & Karthik Varadarajan of SolarCity; Kevin Koy of the UC Berkeley Geospatial Innovation Facility (GIF); Dan Arnold, Lloyd Cibulka, Josiah Johnston, James Nelson, Ciaran Roberts, Michaelangelo Tabone, Alexandra von Meier and Shuyu (Simon) Yang of UC Berkeley; Emma Stewart of Lawrence Berkeley National Lab; and the UCB CITRIS computing cluster for their invaluable assistance.

### 4.3.2 Consolidated Dynamic Pricing of Power System Regulation

**4.3.2.1 Introduction** Power system regulation is a collection of services that adapt in real or near real-time to unanticipated system changes, including automatic generation control, speed droop, and voltage regulation [19, 52, 70, 76]. While it has always been critical to system reliability, the requirement for regulation is increasing due to mismatches between supply and demand induced by renewable variability [44]. In this work, we address how to pay for regulation by developing conceptual pricing methodologies that rigorously incentivize good performance.

Central to the issue is frequency regulation, which refers to actions that maintain the system’s nominal frequency such as fine adjustments in real power generation and consumption. Traditionally, providers have received a capacity payment followed by a market price payment for net energy transacted through frequency regulation. In a recent ruling, the Federal Energy Regulatory Commission deemed this practice inadequate, and mandated that payments be redesigned so that “providers of frequency regulation receive just and reasonable and not unduly discriminatory or preferential rates” [37]. Specifically, the new ruling indicates that the latter payment inadequately incentivizes regulation by failing to capture its cost and quality, and that new ‘pay-for-performance’ frameworks must be developed.

A requirement of the mandate is that regulation resources face a uniform clearing price. ISOs have responded by introducing various forms of mileage payments, in which providers are paid in proportion to the integral of some norm of the regulation signal rather than just its integral [25]. While mileage-based formulations do better account for the actual regulation provided, they do not incorporate dynamic couplings and network effects or the local value and costs of regulation. We contend that uniform prices prevent the inclusion of these features; in the same way that transmission constraints can lead to different locational marginal prices at different locations, each resource’s distinct dynamic constraint and network coupling results in a distinct regulation price.

We present an approach free from these shortcomings that uses the optimal control costate as a regulation price, as originally proposed in [18] and more recently explored in [17, 46, 53]. Indeed, the formulation supports a competitive equilibrium, which is to say that it incentivizes optimal regulation in the same sense that locational marginal prices incentive optimal economic dispatch [69]. Furthermore, by specializing to the linear quadratic regulator, we obtain a pricing policy that is a bilinear function of the control and system state. We comment that historically, this approach as well as optimal feedback control [34, 39] were only of theoretical interest because they required global state information that was not centrally available at the time, a shortcoming now addressed by synchronized phasor measurements [63].

To compute the pricing policy, the system operator must query profit maximizing agents for their costs and values of regulation. As with the dependence of economic dispatch on supply functions and the attendant strategic interaction [43], this constitutes a potential venue for gaming and market power in our approach. To preempt such behaviors, we construct a Vickrey-Clarke-Groves mechanism, which makes truthful revelation of private information a dominant strategy [59].

We subsequently relate our regulation approach to static energy markets following basic cooptimization ideas [80]. Specifically, we represent the infinite horizon LQR as a semidefinite constraint [21, 22], which may then be appended to any economic dispatch constraint set. The resulting formulation tractably incorporates both dynamics and (Gaussian) uncertainty without adding states for future time stages or introducing chance constraints.

Since prices for any (linearized) dynamic service are mechanistically obtainable, the formulation serves as a generic template for consolidating hitherto diverse regulation services such as voltage and frequency regulation at various timescales. As an example, we show that by appropriately defining the control input, the California Independent System Operator's new "flexible ramping product" [24] can be merged with conventional versions of frequency regulation and area control error.

The novel contributions of our work are summarized as follows:

- Application of ideas from [18] to current regulation issues.
- A regulation pricing policy valid at any system state.
- Regulation payment designs, including imbalance fees.
- A Vickrey-Clarke-Groves mechanism to induce honest cost reporting.
- Incorporation of dynamic regulation into static economic dispatch via semidefinite programming.
- Consolidation of dynamic regulation services.
- All formulations are computationally tractable and implementable via in standard software tools.

#### 4.3.2.2 Optimal control

**4.3.2.2.1 Notation** Let  $x_i \in \mathbb{R}^{n_i}$ ,  $u_i \in \mathbb{R}^{m_i}$ , and define  $x \in \mathbb{R}^n$ ,  $n = \sum_i n_i$  and  $u \in \mathbb{R}^m$ ,  $m = \sum_i m_i$  to be  $x_i$  and  $u_i$  respectively stacked column-wise. Let their associated cost matrices be denoted  $Q \in \mathbb{R}^{n \times n}$  and  $R \in \mathbb{R}^{m \times m}$ , with terminal costs denoted by a superscripted  $T$ . The state transition and control input matrices are given by  $A \in \mathbb{R}^{n \times n}$ ,  $B \in \mathbb{R}^{n \times m}$ . The state disturbance is denoted  $w \in \mathbb{R}^n$  with covariance  $W \in \mathbb{R}^{n \times n}$ .

We use single subscripts to denote column blocks of a matrix and double subscripts to denote diagonal blocks; for example,  $B_i \in \mathbb{R}^{n \times m_i}$  and  $B_{ii} \in \mathbb{R}^{n_i \times m_i}$  are the  $i^{th}$  column block and the  $i^{th}$  diagonal block of  $B$ , respectively. We assume that  $Q_{ij} = 0$  and  $R_{ij} = 0$  for all  $i \neq j$ .

To reduce notation and because many of the following expressions are found in standard optimal control textbooks, e.g. [55], we suppress time indexing of dynamic quantities whenever it would be generically  $t$ . We work primarily in continuous time, but note that nearly all formulations are easily converted to discrete time.

**4.3.2.2.2 The classical linear quadratic regulator** Consider the controllable LTI system

$$\dot{x} = Ax + Bu + w, w \sim \mathcal{N}(0, W), x(0) = x_0 \quad (1)$$

and the objective

$$\begin{aligned} J = \min_u \frac{1}{2} \mathbb{E} \sum_i \int_0^T (x_i' Q_{ii} x_i + u_i' R_{ii} u_i) dt \\ + x_i(T) Q_{ii}^T x_i(T). \end{aligned} \quad (2)$$



The optimal control policy is given by

$$u = -R^{-1}B'Px, \quad (3)$$

where  $P \in \mathbb{R}^{n \times n}$  is the solution to the matrix Riccati equation

$$\dot{P} + A'P + PA - PBR^{-1}B'P + Q = 0, P(T) = Q^T.$$

**Lemma 1** *The optimal control policy (3) is attained if each agent  $i$  independently solves*

$$\min_{u_i} \frac{1}{2} u_i' R_{ii} u_i - x' P B_i u_i. \quad (4)$$

**Proof** Differentiating (4) gives the result.

**Remark 1 (Lagrangian duality)** *Lemma 1 can be straightforwardly derived using Lagrangian duality, paralleling standard locational marginal pricing [69] and the development of [18].*

The objective (2) may be written

$$\text{Tr}P(0)m_0m_0' + \text{Tr}P(0)X_0 + \text{Tr} \int_0^T PW dt,$$

where  $m_0$  and  $X_0$  are respectively the state's initial mean and covariance. We will require similar expressions for generic objectives and dynamics. Suppose that the state evolves according to  $\Phi \in \mathbb{R}^{n \times n}$ . Let the state's mean and covariance be  $m \in \mathbb{R}^n$  and  $X \in \mathbb{R}^{n \times n}$ , which evolve from zero initial conditions under  $\Phi$  as

$$\begin{aligned} \dot{m} &= \Phi m, m(0) = 0 \\ \dot{X} &= \Phi X + X\Phi' + W, X(0) = 0. \end{aligned}$$

**Lemma 2 ([55])** *Define*

$$\begin{aligned} \Lambda(\Phi, \Psi) &= e^{\Phi'(T-t)} \Psi^T e^{\Phi(T-t)} \\ &\quad + \int_t^T e^{\Phi'(T-\tau)} \Psi e^{\Phi(T-\tau)} d\tau, \end{aligned}$$

*which is the solution to the Lyapunov differential equation*

$$\dot{\Lambda} + \Lambda\Phi + \Phi'\Lambda + \Psi = 0, \Lambda(T) = \Psi^T.$$

*Then*

$$\frac{1}{2} \mathbb{E} \left[ \int_t^T x' \Psi x dt + x(T)' \Psi^T x(T) \right] = \text{Tr} \int_0^T \Lambda(\Phi, \Psi) W dt.$$

#### 4.3.2.3 Economic interpretation

**4.3.2.3.1 Modeling assumptions** In this section, we will use (1) to model the physics of a power system, linearized about an operating point. For example, the  $A$  and  $B$  matrices could represent system voltage or frequency dynamics, in which cases  $x$  is voltage magnitude or voltage angle and  $u$  could be reactive or real power, respectively. The latter scenario is examined in Section 4.3.2.5.2. Per usual in linear quadratic regulation, the state and control vectors  $x$  and  $u$  are unconstrained [55]. We comment that the linearization precludes application to energy markets, because energy markets do not track an operating point subject to disturbances. Rather, the operating point could be the outcome of an energy market, which we consider in Section 4.3.2.5.1.

The objective (2) represents penalties associated with deviating from the operating point and taking control actions, for example factory production losses, battery life degradation, or generator fuel consumption. Details such as voltage and line flow limits can also be incorporated as soft constraints via penalty terms in the objective. While this is a departure from current power system regulation procedures, in which physical quantities like voltage are required to stay within prescribed limits, it is consistent with the Federal Energy Regulatory Commission’s recent ‘pay-for-performance’ mandate [37].

In this section, we assume that the system operator exactly knows the matrices  $A$ ,  $B$ ,  $Q$ , and  $R$ , which completely parametrize our framework. In Section 4.3.2.4.2, we consider the more realistic case that  $Q$  and  $R$  are revealed by the market participants, and address dishonest reporting with a Vickrey-Clarke-Groves mechanism.

The matrix  $Q$  can be used to tune the relative importance of system deviations at different locations in the network. For example, bus  $i$  might serve a collection of customers that require tight frequency performance but can tolerate large voltage fluctuations, in which case the entries of  $Q$  for a particular bus would be analogous to the “satisfaction index” discussed in [18].

Moreover, for generators with (approximately) quadratic heat rate curves,  $R_{ii}$  can be directly interpreted in terms of economic costs [70, 76]. We note that the LQR framework also admits linear terms on the control variables in the objective function [55], which would improve the fidelity of the heat rate model.  $R_{ii}$  could also be used to model the flexibility or willingness of the demand side to respond to a need for control actions.

Unlike  $R$ , it may not always be possible to define the  $Q$  parameters directly in terms of real economic costs. However the parameters can nonetheless be tuned to achieve acceptable outcomes, for example voltage and frequency excursions that are usually within a desired range.  $Q$  values could be set, for example, throughout the network by the power system operator, in collaboration with or with input from individual customers or load serving entities.

The Gaussian assumption is reasonable for uncertainty from load and renewable aggregations [19, 76], but cannot accommodate contingencies, a major source of power system disturbances. We remark that contingency reserves are separate from regulation [70] and not addressed by our framework, but that incorporating contingencies is an important future direction.

**4.3.2.3.2 Pricing regulation** Similar to [18], agent  $i$ ’s control  $u_i$  is economically related to this cost through Lemma 1, which motivates the following definition:

**Definition 1 (Pricing policy)** *The quantity  $x'PB_i$  is agent  $i$ ’s locational marginal regulation price.*

The most important property of this price is that it supports a competitive equilibrium: it is most profitable for agent  $i$  to choose the  $u_i$  that is centrally optimal, which happens to be the LQR control policy. Moreover, the fact that it is a policy adds essential versatility: since it is valid at any system state, it is well-suited to unanticipated changes. It also inherits all standard LQR properties, including certainty equivalence and the separation principle if the state is estimated using a Kalman filter. As discussed later, this feature enables uncertainty to be tractably incorporated.

**Remark 2 (Dynamic pricing versus ex-post settlement)** *While prices are often used in other contexts to actively moderate supply and demand [56], here we are not exclusively suggesting that regulation be provided in response to a price, which may be unrealistic at the time scales frequency and voltage excursions occur over. Rather,  $x'PB_i$  could also determine a regulation provider's ex-post settlement, so that they ultimately receive*

$$\int_0^T x'PB_i u_i dt, \quad (5)$$

where  $x$  and  $u$  are the actual state and control trajectories in  $[0, T]$ . In this case,  $u$  could be a centrally commanded action such as automatic generation control.

**Remark 3 (Non-LQR control)** *In (5),  $u_i$  is not necessarily the LQR control policy (3), but any control provided by agent  $i$ ; in other words, the LQR control policy is the equilibrium supported by these prices, but need not be the actual control.*

**4.3.2.3.3 The value of regulation** We now use Lemma 2 to evaluate various costs associated with regulation. We consider both realized and expected quantities, in the case of the latter assuming that the LQR policy (3) has been employed by each agent. We only present instantaneous quantities, from which ex-post quantities can be obtained through integration.

Define

$$\begin{aligned} F &= A - BR^{-1}B'P \\ U^i &= PB_i R_{ii}^{-1} B_i' P \end{aligned}$$

Let  $U = \sum_i U^i$  and  $Q_{ii}^0 \in \mathbb{R}^{n \times n}$  be equal to  $Q_{ii}$  on the  $i^{th}$  block of its main diagonal and be zero elsewhere. To simplify our exposition, we also assume that the system begins exactly at its equilibrium point so that  $m_0 = 0$  and  $X_0 = 0$ .

The first quantity of interest is the expected instantaneous regulation payment to agent  $i$ :

$$\frac{1}{2} \mathbb{E} x' U^i x.$$

Lemma 2 tells us that it is given by the following definition:

**Definition 2 (Cost of regulation)** *The expected and actual regulation payments to agent  $i$  are*

$$\begin{aligned} \rho_i &= \text{Tr} \Lambda(F, U^i) \text{ and} \\ \hat{\rho}_i &= x' P B_i u_i. \end{aligned}$$

We proceed similarly for other quantities of interest.

**Definition 3 (Benefit of regulation)** *The expected and actual regulation induced reduction in agent  $i$ 's operating cost are*

$$\nu_i = \text{Tr}(\Lambda(A, Q_{ii}^0) - \Lambda(F, Q_{ii}^0)) \text{ and} \quad (6)$$

$$\hat{\nu}_i = \text{Tr}(\Lambda(A, Q_{ii}^0) W - x' Q_{ii}^0 x). \quad (7)$$

Now consider the contribution of disturbances from agent  $i$  to the cost of regulation. The total cost of regulation may be written

$$\begin{aligned} \text{Tr} \Lambda(F, U) W &= \sum_i \text{Tr} \Lambda(F, U)_{ii} W_{ii} \\ &\quad + 2 \sum_{j \neq i} \text{Tr} \Lambda(F, U)_{ij} W_{ij} \end{aligned}$$

The first summation may be unambiguously divided index-wise among the agents, but a sharing rule must be defined for the latter. A naive scheme is to attribute  $\text{Tr} \Lambda(F, U)_{ij} W_{ij}$  each to agents  $i$  and  $j$ , yielding the net contribution  $\text{Tr} \Lambda(F, U)'_i W_i$ ; this however may disproportionately penalize small disturbance sources for being correlated with larger disturbance sources.

We instead suggest the following measure:

**Definition 4 (Expected cost of disturbances)** *The proportionally fair expected cost of disturbances from agent  $i$  is*

$$\begin{aligned} \mu_i &= \text{Tr} \Lambda(F, U)_{ii} W_{ii} \\ &\quad + 2 \sum_{j \neq i} \frac{\text{Tr} W_{ii}}{\text{Tr} W_{ii} + \text{Tr} W_{jj}} \text{Tr} \Lambda(F, U)_{ij} W_{ij}. \end{aligned} \quad (8)$$

Because the disturbance vector  $w$  does not appear explicitly in (2) and because the actual trajectory may not correspond to the idealized system dynamics, we cannot give an exact expression for  $\hat{\mu}_i$ , but rather use  $\mu_i$  to design statistically correct payments. If we assume a symmetric imbalance fee  $\gamma_i$ , we should have

$$\int_0^T \gamma_i \mathbb{E} |w_i| dt = \int_0^T \mu_i dt.$$

The imbalance fee may either be time-varying or constant, which respectively yield

$$\gamma_i = \frac{\mu_i}{\mathbb{E} |w_i|} \quad \text{and} \quad \gamma_i = \frac{\int_0^T \mu_i dt}{\int_0^T \mathbb{E} |w_i| dt},$$

where

$$\mathbb{E} |w_i| = 2 \sqrt{\frac{2W_{ii}}{\pi}}.$$

**Definition 5 (Imbalance cost)** *The actual imbalance-wise cost of disturbances from agent  $i$  is*

$$\hat{\mu}_i = \gamma_i |w_i|.$$

**Remark 4 (Agent to agent costs)** *By replacing  $F$  with  $A - B_j R_{jj}^{-1} B_j' P$  in (6) and (8), we can isolate the contribution of agent  $j$ 's regulation to  $\nu_i$  and  $\mu_i$ , which may be useful for designing bilateral contracts.*

**4.3.2.4 Designing transactions** We now discuss how (5) should be financed, given that the system operator must be revenue neutral in some sense. An inherent ambiguity arises in choosing who should make the payment: in the absence of disturbances, no regulation and hence payment is necessary. Contrarily, if no consumer valued regulation, i.e. if  $Q = 0$ , regulation is also unnecessary. Should then disturbance sources or the beneficiaries of regulation pay for (5)? We do not address this question directly, but rather calculate quantities that would be involved in designing such payments.

**4.3.2.4.1 Payments** We respectively consider disturbance and regulation payments of the form

$$\alpha \phi_i^\mu(\hat{\mu}) \quad \text{and} \quad (1 - \alpha) \phi_i^\nu(\hat{\nu}),$$

that satisfy

$$\mathbb{E} \sum_i \phi_i^\mu(\hat{\mu}) = \mathbb{E} \sum_i \phi_i^\nu(\hat{\nu}) = \chi,$$

where  $\chi = \sum_i \rho_i$  and  $\alpha \in [0, 1]$ .

**Definition 6** *Agent  $i$ 's net instantaneous cost is*

$$\begin{aligned} \hat{J}_i = & \frac{1}{2} (x_i' Q_{ii} x_i + u_i' R_{ii} u_i) - x_i' P B_i u_i \\ & + \alpha \phi_i^\mu(\hat{\mu}) + (1 - \alpha) \phi_i^\nu(\hat{\nu}). \end{aligned} \tag{9}$$

We now dissect this payment.

- The parameter  $\alpha$  determines what portions of (5) should be paid by the disturbance sources versus the regulation beneficiaries; determining the correct value of  $\alpha$  is outside of our current scope.
- We have not specified a particular payment structure because no single format may be universally appropriate. For example, if wind farms are to be directly accountable for their intermittency, those producing larger disturbances should pay more, and the proportionally fair payment  $\phi_i^\mu(\hat{\mu}) = \hat{\mu}$  is appropriate. On the other hand, it could be difficult and contentious to audit household demand variability, in which case the constant  $\phi_i^\mu(\hat{\mu}) = \chi/n$  is more appropriate. Similar analogies hold for regulation beneficiaries, where  $\phi_i^\nu(\hat{\nu}) = \chi \hat{\nu} / \sum_i \nu_i$  and  $\phi_i^\nu(\hat{\nu}) = \chi/n$  are the proportionally fair and constant payments, respectively.
- Since neither of the latter two terms in (9) explicitly depend on  $u_i$ , the LQR competitive equilibrium remains as in (4). This is clear upon differentiating  $\hat{J}_i$  with respect to  $u_i$ .
- In expectation, the system operator is revenue neutral. This may not be true for a particular realization of uncertainty, placing the onus on the system operator to absorb temporary financial imbalances. However, if uncertainties have been accurately modeled, it can be expected that the system operator will become revenue neutral on average over long time periods.

**4.3.2.4.2 Preempting strategic behavior** The matrices  $Q_{ii}$  and  $R_{ii}$  depend largely on information provided by the profit-maximizing agents, and thus comprise a potential venue for gaming. In this section, we construct a Vickrey-Clarke-Groves (VCG) mechanism to induce truthful revelation of these values [59]. The VCG mechanism specifies an additional upfront tax as a function of the reported matrices  $Q_{ii}$  and  $R_{ii}$ .

This tax makes honesty a dominant strategy, which is to say that honesty is optimal for each agent regardless of the others' actions, e.g. if some agents collude in reporting their cost parameters.

We assume that  $Q_{ii}$  and  $R_{ii}$  are reported at  $t = 0$ . Define

$$\begin{aligned} J_i &= \mathbb{E} \int_0^T \hat{J}_i dt \\ &= \int_0^T \text{Tr} \Lambda (F, Q_{ii}^0 + U^i) W - \rho_i \\ &\quad + \mathbb{E} [\alpha \phi_i^\mu(\hat{\mu}) + (1 - \alpha) \phi_i^\nu(\hat{\nu})] dt, \end{aligned}$$

and note that  $\sum_i J_i = J$ .

Define  $R_{-i}$  to be  $R$  with the  $i^{th}$  row and column block removed, and let  $P_{-i} \in \mathbb{R}^{n \times n}$  be the solution to the Riccati equation

$$\dot{P}_{-i} + A' P_{-i} + P_{-i} A - P_{-i} B_{-i} R_{-i}^{-1} B_{-i}' P_{-i} + Q - Q_{ii}^0 = 0,$$

where

$$\begin{aligned} B_{-j} &= [B_1, \dots, B_{j-1}, B_{j+1}, \dots, B_n] \\ u_{-j} &= [u_1, \dots, u_{j-1}, u_{j+1}, \dots, u_n]' \\ F_{-i} &= A - B_{-i} (R_{-i})^{-1} (B_{-i})' P_{-i} \end{aligned}$$

Consider the LTI system

$$\dot{x} = Ax + B_{-j} u_{-j} + w,$$

and the objective

$$\begin{aligned} J_{-i} &= \min_{u_{-i}} \mathbb{E} \frac{1}{2} \sum_{j=-i} \int_0^T (x_j' Q_{jj} x_j + u_j' R_{jj} u_j) dt \\ &= \int_0^T \text{Tr} \Lambda (F_{-i}, Q - Q_{ii}^0 + U - U^i) W dt. \end{aligned}$$

**Definition 7** Let  $(Q_{ii}, R_{ii})$ ,  $i = 1, \dots, n$  be the pairs cost matrices reported by each agent. The VCG tax for agent  $i$  is given by

$$\beta_i = \sum_{j=-i} J_j - J_{-i}. \quad (10)$$

**Lemma 3** Given the tax function of (10), truthful reporting of the matrices  $Q_{ii}, R_{ii}$  is a dominant strategy for each agent  $i$ .

**Proof** Firstly, observe that the term  $J_{-i}$  does not depend on the matrices reported by agent  $i$ . Now, suppose agent  $i$  reports  $(\tilde{Q}_{ii}, \tilde{R}_{ii})$  that are different from the true  $(Q_{ii}, R_{ii})$ . Then, the locational marginal regulation prices based on the mis-reported values will result in a (socially) sub-optimal control policy and a net social cost of  $\tilde{J} = \sum_k \tilde{J}_k \geq J$ . Further, the tax for agent  $i$  would be  $\sum_{j=-i} \tilde{J}_j - J_{-i}$ . Agent  $i$ 's net cost,  $\tilde{J}_i + \beta_i$ , would then be

$$\begin{aligned} \tilde{J}_i + \sum_{j=-i} \tilde{J}_j - J_{-i} &= \tilde{J} - J_{-i} \\ &\geq J - J_{-i}. \end{aligned} \quad (11)$$

On the other hand, if  $(\tilde{Q}_{ii}, \tilde{R}_{ii}) = (Q_{ii}, R_{ii})$ , then the inequality in (11) would become an equality. Therefore, truthful reporting is a dominant strategy for agent  $i$ .

**Remark 5**  $Q_{ii}$  may be much harder to quantify than  $R_{ii}$ , and moreover may not be accurately known to agent  $i$ . As discussed in Section 4.3.2.3.1, it is viable that only  $R_{ii}$  be reported by agent  $i$ , and that  $Q_{ii}$  be set by a regulating body such as an ISO. In this case,  $Q_{ii}$  is not a revealed value, and so the VCG mechanism only functions to induce honest reporting of  $R_{ii}$ .

**Remark 6** The VCG taxes may not be budget-balanced. That is,  $\sum \beta_i \neq 0$ . One way of reducing the budget imbalance is to modify the tax to:

$$\beta_i = \sum_{j=-i} J_j - sJ_{-i}$$

where  $s$  is a scaling factor that can be appropriately chosen to reduce the budget imbalance. Even with this modified tax, truthful reporting is still a dominant strategy for each agent. The impossibility results from [48] suggest that an incentive compatible mechanism with no budget imbalance may not be possible.

#### 4.3.2.5 Application to power systems

**4.3.2.5.1 Primary and regulation energy cooptimization** The price in Definition 1 is a dynamic analog to standard locational marginal pricing. Since both are based on fundamental duality principles, we can combine them into a single framework, which is interpretable as jointly pricing energy and regulation. Indeed, the linearization of their dynamics and costs in (1) and (2) depend on the outcomes of static energy markets, justifying co-optimized approaches as are currently practiced in some primary and regulation markets [80]. We now put forth such an approach incorporating regulation prices from an LQR. Note that our intention is not to construct a practical co-optimization model, which is outside of our present scope, but to demonstrate an analytical mechanism by which one could do so.

Suppose that the objective is to minimize the cost of power plus regulation, where real power from bus  $i$  is denoted  $p_i$  and its cost by  $c_i(p_i)$ . Suppose further that the linearized dynamics  $A$  and  $B$  are independent of  $p$ . Denote the feasible set of power injections by  $\Delta$ , and assume that the cost of regulation is a function of real power  $R(p)$  such that  $R(p)^{-1}$  is affine in  $p$ . For example, if the cost of regulation from bus  $i$  increases with its base real power production,  $R(p)_{ii}^{-1} = R_{ii}^1 p_i + R_{ii}^2$  with  $R_{ii}^1, R_{ii}^2 > 0$ .

It is well known that for fixed  $p$ , the below semidefinite program reproduces the infinite horizon LQR with  $P = S^{-1}$  [21, 22]; note that  $Y$  is an auxiliary variable, and that additional constraints for  $S^{-1}$  are omitted as they are standard. Declaring  $p$  a variable yields an optimization over power injections given that regulation will also be provided from the same buses, the cost of which is coupled to  $p$  through  $R(p)$ ; in other words, it is a quasistatic optimal power flow that accounts for dynamic regulation costs.

$$\begin{aligned}
& \min_{p, S, X, Y} \quad \sum_i c_i(p_i) + \text{Tr} S^{-1} W \\
& \text{s.t.} \quad p \in \Delta, \quad S \succeq 0 \\
& \quad \begin{bmatrix} SA' + AS & S & Y' \\ +BY + Y'B' & -Q^{-1} & 0 \\ S & 0 & -R(p)^{-1} \\ Y & & \end{bmatrix} \preceq 0
\end{aligned}$$

We'd like to construct prices which recover the first objective term and the portion of the second corresponding to regulation costs ( $u' Ru$ ). Because the objective is separable in  $p$  and  $S$ , we can find separate primary energy prices that influence the first term and regulation prices that influence the second as follows:

- *Primary energy prices:* All multipliers associated with the constraint  $p \in \Delta$  and those associated with  $R(p)^{-1}$  in the semidefinite constraint influence the first objective term. These multipliers completely account for the first objective term, and can be used as cooptimized primary energy prices.
- *Regulation prices:* All remaining multipliers are thus sensitivities of the second objective term, and therefore its costs are recovered via regulation payments  $x' S^{-1} B_i u_i$  as in Section 4.3.2.3.2.

Note that semidefinite programming duality is essentially the same as the linear case; the reader is referred to [22] for in depth discussion.

We make the following observations about this approach.

- It tractably incorporates uncertainty and its dynamic effects, encapsulated in  $W$ , into optimal power flow routines.
- Since the new regulation constraints are convex, convexity is retained if they are incorporated in DC (linear) optimal power flow as well as recent convex relaxations [16, 22, 47].
- The dependence of bus  $i$ 's control cost on  $p_i$  is more realistically but complexly captured by the LQR control objective  $R_{ii} (p_i + p'_i)^2$ , where  $p'_i$  is the LQR control input. This is an extension of LQR known as a 'tracking problem', and is handled by standard formulations [55].

**4.3.2.5.2 Example: frequency regulation** Renewable variability has increased need for active power regulation resources that can make large adjustments to their output levels over short time intervals. In recognition, the California Independent System Operator has introduced a "flexible ramping product" [24], in which ramping energy and shadow prices are determined via optimal power flow [78]. In this example, we consolidate conventional frequency regulation, flexible ramping, and area control error in one formulation. We formulate this problem in both continuous and discrete time to highlight differences between the two.

**Continuous time** Define  $\omega_i$ ,  $\theta_i$ , and  $u_i$  to be the nominal frequency deviation, voltage angle, and real power injection at bus  $i$ . We regard  $u_i$  as a state, and define the control input  $v_i$  to be the derivative of  $u_i$ .



Let  $I, D, H$  respectively be the identity matrix (appropriately sized), a positive, diagonal damping matrix, and a diagonal matrix of rotor inertias. Define

$$L_{ij} = \begin{cases} b_{ij} & i \neq j \\ -\sum_k b_{ik} & i = j \end{cases},$$

where  $b_{ij} \in \mathbb{R}^+$  is the inductance of line  $ij$ . The linearized system dynamics are given by

$$\begin{bmatrix} \dot{\omega} \\ \dot{\theta} \\ \dot{u} \end{bmatrix} = \begin{bmatrix} -H^{-1}D & H^{-1}L & H^{-1} \\ \omega_0 I & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \omega \\ \theta \\ u \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ I \end{bmatrix} v + \begin{bmatrix} w \\ 0 \\ 0 \end{bmatrix}, w \sim \mathcal{N}(0, W).$$

One can straightforwardly verify that this system is controllable. Consider the infinite horizon objective

$$\begin{aligned} \min_v \lim_{T \rightarrow \infty} \frac{1}{2T} \mathbb{E} \int_0^T \sum_i (Q_{ii}^\omega \omega_i^2 + R_{ii}^u u_i^2 + R_{ii}^v v_i^2) \\ + \sum_{(i,j) \in \mathcal{A}} Q_{ij}^\theta b_{ij}^2 (\theta_i - \theta_j)^2 dt, \end{aligned} \quad (12)$$

where  $\mathcal{A}$  is some subset of lines. The first and second terms in the first sum are standard and penalize frequency deviations and regulation power injections, and the third is a ramping cost that penalizes rapid changes in control input. The second summation accounts for undesired power flows, e.g. if  $\mathcal{A}$  is set of inter-area lines, it could serve as a surrogate for area control error.

If we naively apply the pricing scheme of Lemma 1, agent  $i$ 's regulation payment is  $x' P B_i v_i$ , which, if all agents apply LQR, is equal in expectation to the third term of (12). Agent  $i$  is thus being paid for their nominal control action  $v_i$ , but not the actual regulation they provide,  $v_i$  and  $u_i$ . The inconsistency stems from the fact that  $v_i$  has an infinitesimal effect on  $u_i$  at time  $t$ , which admits the unrealistic interpretation that agent  $i$  is unaware of  $v_i$ 's influence on  $u_i$ . A solution consistent with this assumption is to simply also pay agent  $i$   $u_i' R_{ii} u_i$  at each time. Indeed, since  $u_i$  is not a control variable, this retains all of the desirable properties of the locational marginal regulation price, albeit in a narrow technical sense.

**Discrete time** By switching to discrete time, we can consistently model both the power injection and ramping costs with the same control input. We now regard the real power injection  $u$  as the control and define the new state  $x$  to be  $u$  at the previous time step. Letting  $z$  and  $\delta$  respectively be the forward shift operator and time step, the discrete time dynamics are given by

$$\begin{aligned} z \begin{bmatrix} \omega \\ \theta \\ x \end{bmatrix} &= \left( I + \delta \begin{bmatrix} -H^{-1}D & H^{-1}L & 0 \\ \omega_0 I & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right) \begin{bmatrix} \omega \\ \theta \\ x \end{bmatrix} \\ &+ \delta \begin{bmatrix} I \\ 0 \\ I \end{bmatrix} u + \begin{bmatrix} w \\ 0 \\ 0 \end{bmatrix}, w \sim \mathcal{N}(0, \sqrt{\delta} W). \end{aligned}$$

The new objective is given by

$$\min_v \lim_{T \rightarrow \infty} \frac{\delta}{2T} \mathbb{E} \sum_{t=0}^T \sum_i (Q_{ii}^\omega \omega_i^2 + R_{ii}^u u_i^2 + R_{ii}^v (u_i - x_i)^2) + \sum_{(i,j) \in \mathcal{A}} Q_{ij}^\theta b_{ij}^2 (\theta_i - \theta_j)^2.$$

Note that the the control  $u_i$  now appears finitely in both the power injection and ramping terms. The new third term is also convex, and can be dealt with using standard LQR techniques [55].

**Remark 7 (VCG convention)** *There are different ways of defining agent  $i$ 's participation for computing quantities in Section 4.3.2.4.2. For example one could model the absence of agent  $i$  by setting its control and costs to zero, but leaving  $A$  unchanged, in turn making the system harder to control for the remaining agents. Here we instead model agent  $i$ 's absence as disconnection from the system by Kron reducing the inductance matrix  $L$  [66] and removing the appropriate rows and columns from the remaining matrices.*

**Numerical results** We now examine some characteristics of our approach in this example. We consider a discrete time, nine-bus system with the dynamics of the previous section, in which the control input  $u$  and corresponding price are from the infinite horizon LQR solution. Parameter values are given by  $\delta = 0.01$ ,  $\omega_0 = 120\pi$ ,  $H = 8I$ ,  $D = 5I$ ,  $Q = 10 \times \text{diag}[1 \ 2 \ 3 \ 1 \ 2 \ 3 \ 1 \ 2 \ 3]$ ,  $R^u = 10 \times \text{diag}[3 \ 2 \ 1 \ 3 \ 2 \ 1 \ 3 \ 2 \ 1]$ ,  $R^v = 10 \times \text{diag}[1 \ \cdots \ 1]$ , and  $W = \delta^2 \text{diag}[1 \ 1 \ 1 \ 2 \ 2 \ 2 \ 3 \ 3 \ 3]$ . Line inductances are taken from the nine-bus test case in MATPOWER [81].

Fig. 9 shows the power injection control input and the corresponding price at buses two and seven for a sample disturbance trajectory. As is easily shown analytically, the two signals at a particular bus have the same sign and thus a positive payment if each bus employs LQR, but the price and optimal control may be of different signs at different buses, as seen just before 0.4 seconds.

Fig. 10 shows the expected average regulation payment and VCG tax for each bus on top and the expected average disturbance cost and imbalance fee on bottom. The VCG tax is smaller than the expected payment, but of the same order; this is among many known issues with VCG mechanisms, justifying further work in more practical schemes [65]. As expected, the imbalance fee is proportional to the square root of the expected average cost of disturbances, which is desirable because it reduces the spread of imbalance fees across buses.

**4.3.2.6 Conclusions and future work** Regulation payments must necessarily be based on dynamic models to correctly incorporate the fundamentally dynamic nature of power system regulation and incentive the corresponding desirable actions. In this work, we present regulation pricing mechanisms based on the optimal control-based pricing ideas of [18], which are well-suited to current regulation issues [24, 37]. The approach shares standard locational marginal pricing's theoretical underpinnings, and likewise supports a competitive equilibrium in which linear quadratic optimal regulation is incentivized. By working within an LQR framework and employing tools from mechanism design and convex optimization, we have made the original approach considerably more versatile and practical.

Our main conclusions are as follows.

- Over-differentiation of regulation services can lead to inefficient utilization and markets, particularly when the services are overlapping. Our approach can generically consolidate regulation services

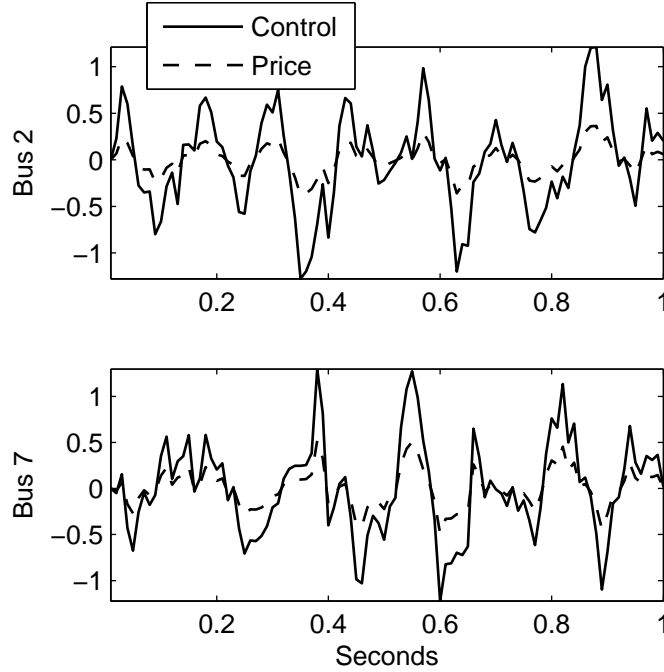


Figure 9: Control input, price, and payment at buses two and seven.

with linear(izable) dynamic descriptions, incentivizing more efficient, joint utilization and markets.

- The analytical properties of LQR enable efficient, first principles computation of many economic quantities such as imbalance fees and regulation benefits. In this regard, our approach constitutes a practical tool for design and analysis in dynamic scenarios, similar to the use of DC and decoupled optimal power flow in quasistatic models.
- The most computationally intensive aspect of our approach is solving a matrix Riccati equation. Powerful, standardized algorithms exist for this computation, imparting our approach with the scalability necessary for practical market scenarios.
- Our approach is consistent with the shift from operating limits to ‘pay-for-performance’ [37]. This necessarily induces reliance on private information from potentially dishonest agents in the same sense that economic dispatch relies on supply functions. Algorithmic game theory and mechanism design offer tools such as the VCG mechanism for inducing honest participation in such scenarios.

We now discuss some current limitations and future directions.

- Dead-zones and hard constraints, e.g. on ramping, are important and realistic physical details, which could be incorporated using hybrid and constrained linear quadratic regulation, but at great computational expense if a policy is desired [20].
- Contingencies are a major source of disturbances which should be incorporated, but cannot be modeled as Gaussian noise.
- While the VCG mechanism makes honesty a dominant strategy, the formulation does not provide a means of testing for manipulation that has occurred.

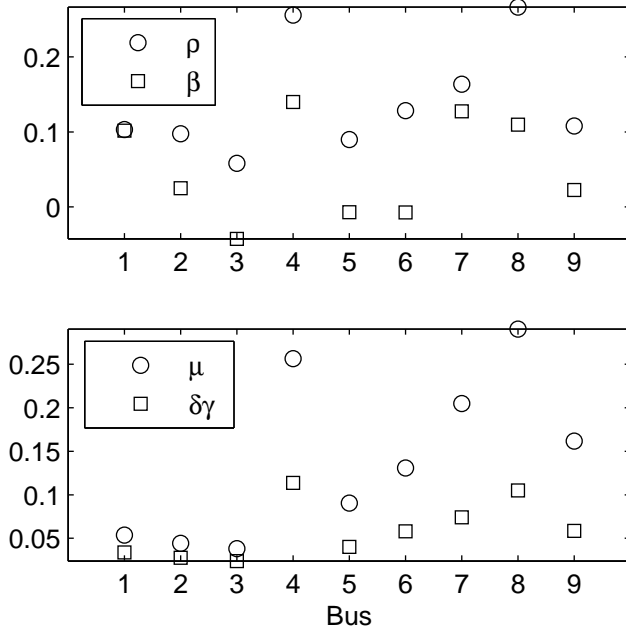


Figure 10: Top: Bus-wise expected payments  $\rho$  and VCG taxes  $\beta$ , bottom: expected costs of disturbances  $\mu$  and time step times imbalance fee  $\delta\gamma$ .

- Game theoretic analysis must be performed to determine strategic outcomes when agents anticipate the effect of their control action on future states and prices.

**Acknowledgement** We thank Professors Pravin Varaiya, Shmuel Oren, and Andrew Packard for helpful discussion.

### 4.3.3 Price and capacity competition in zero-mean storage and demand response markets

**4.3.3.1 Introduction** Energy storage and demand response are widely considered promising solutions to renewable variability [27, 28]. Independent System Operators (ISOs) and Regional Transmission Operators (RTOs) have begun to propose new frameworks to encourage such ‘non-generator resources’ (NGRs) to participate in regulation markets [26, 32]. These new frameworks must address the fact that NGRs have different capabilities and constraints than power plants that have traditionally provided these services. Specifically, NGRs have small-to-no ramp constraints but they do have strict energy capacity constraints, an implication of which is that the total amount of energy transacted must be ‘zero-mean’ over time.

Traditional regulation and spinning reserve markets are run as capacity auctions [13], in which power capacity is set aside for a premium or capacity payment in a day-ahead market, and then purchased at the real-time market price as needed. Characteristics of generators make capacity payments an appropriate format; specifically, generator opportunity costs and unit commitment constraints necessitate day-ahead planning [70]. Moreover, generator reserves provided at one time are more or less independent from those subsequently needed because fuel is generally available in sufficient quantity. On the other hand, NGRs do not have unit commitment constraints, opportunity costs are substantially lower because they cannot sell energy without first buying it in the same form, and the capacity of storage depends on prior decisions. As such, it is worth considering alternative market formats.

In this work, we analyze two formats for NGR markets:

1. Ex-ante capacity payments as used for conventional reserves, and
2. Imbalance fees for providing *or consuming* energy.

While non-traditional, imbalance fees are conceptually similar to mileage payments, versions of which will soon see inclusion in regulation payment mechanisms [25] in response to a recent Federal Energy Regulatory Commission (FERC) ruling [37] mandating that regulation providers be paid more fairly.

We analyze each format using game theory, which has seen extensive application in analyzing conventional power markets. Generator competition is often modeled using supply function equilibrium [43, 50], because bids submitted by generators are often actual supply functions, which reflect nonlinear fuel curves and other generation costs. For NGRs, however, the cost/value of energy is more or less linear with quantity due to negligible marginal costs.

In lieu of these considerations, we model NGR markets using Bertrand-Edgeworth competition [33], in which energy storage is patronized up to capacity on the basis of price. This framework was applied to electricity market design in [36]. Bertrand-Edgeworth competition has been widely applied within scenarios in which firms first compete on capacity and then price, beginning with [51] in which such scenarios were shown to produce Cournot outcomes. Since then, however, limited attention has been devoted to uncertainty, cf. [15, 54, 64]. In this work, we obtain new results for the case that the object of competition is random, i.e. the energy associated with providing frequency regulation. We build primarily on the work of [12], which is also concerned with capacity followed by price competition.

Ultimately, we obtain a comparison between capacity and mileage payments. Since future markets will likely contain both [25], our work represents an extreme case analysis, which identifies the characteristics each format imparts. For concision, all proofs are omitted.

#### 4.3.3.2 Setup

**4.3.3.2.1 Non-generating resources** In this work, NGR refers to any energy storage device (e.g., battery, pumped hydro, compressed air, flywheel, capacitor) or demand response resource capable of shifting energy usage in time (e.g., thermal energy storage in buildings and appliances, industrial process shifting). We do not consider demand response that reduces total energy consumption, as this results in an a decrease in service to the electricity customer.

All NGRs have both energy and power capacity constraints. Ignoring energy losses via static leakage and charge/discharge efficiencies, energy capacity constraints require that an NGR follow a signal that is zero mean with respect to its nominal operating point. Zero-mean real-time markets are possible via ISO procurement of the expected value of net load in both the day-ahead and day-of markets, instead of that minus a bias as is commonly done today. The time-scale over which the real-time market is zero mean is dictated by the scheduling of the proceeding market. Alternatively, zero-mean markets could be achieved by using other markets to manage the state of charge of NGRs, periodically resetting them to their nominal operating points, for example, as proposed in California [26] and Texas [32].

NGRs can respond to signals very quickly, especially as compared to most traditional generators such as coal, nuclear, and steam turbines which can only ramp 1–5%/min, while gas turbines and diesel generators respond faster, ramping 5–40%/min [75]. In contrast, a pumped hydro NGR in Wales, UK can ramp to full output in less than 16 seconds [31]. Batteries can likely respond even faster. For demand response, both telemetry and building automation hardware/software dictate the speed of the response. In commercial buildings using temperature set point control for demand response, demonstration projects have shown full responses within 3–8 minutes [58]. Residential air conditioners can fully respond within 80–90 seconds [35, 49]. Since NGRs are fast-acting and to simplify our analysis, we ignore power capacity constraints in our formulation.

NGRs have low marginal costs, dictated by operations/maintenance costs and energy losses. Therefore, NGR marginal costs are often assumed to be zero. Here, we assume the marginal cost of an NGR is small and constant with respect to power, up to the actual or effective power rating of the NGR. This assumption is consistent with the storage literature [62] and the demand response literature [14, 73].

**4.3.3.2.2 Competitive framework** In this section, we outline the two market formats of interest and their relevant quantities. We distinguish between two time periods: the forward market (FM), in which scheduling and capacity payments are made, and the real-time market (RTM), in which energy transactions and payments are made; note that by FM we simply mean ahead of real-time, and so this could also mean hour ahead. We will henceforth exclusively refer to these periods as the FM and RTM.

Market participants are the system operator and the NGRs. The system operator wishes to procure a total of  $\Delta$  reserves to address zero-mean energy fluctuations, e.g.  $\Delta$  might be a multiple of the standard deviation of the load forecast uncertainty [42].

We would like to compare procurement of NGR capacity in the FM versus the RTM via generic market models. The two scenarios, which we denote **M1** and **M2**, are described in Tables 3 and 4. In the former, a deterministic capacity  $C$  is competed for through premiums, while in the latter, a random energy demand  $D$  is competed for prior to realization through imbalance fees, as summarized in Table 5. Note that

imbalance fees are always positive, and represent either a discount (negative demand) or bonus (positive demand) additional to the market price. Because the market price portion of each energy transaction is identical in both formats, we do not include this quantity in our models.

In game theoretic terms, firm  $i$ 's strategy is composed of its committed capacity,  $S_i \leq \mathbb{S}_i$ , and price (premium or imbalance fee),  $p_i$ . The market operator in both formats is modeled as a Stackelberg leader who desires a capacity  $0 \leq C \leq \Delta$  from the NGRs.

Table 3: Sequence of events in **M1**

Stage	Time	Description	Competition
1	FM	The system operator requests capacity $C \leq \Delta$ from NGRs and purchases $\Delta - C$ generator reserves.	–
2	FM	Each firm $i$ specifies a maximum capacity $S_i \leq \mathbb{S}_i$ to the system operator.	Capacity
3	FM	Each firm $i$ offers a capacity premium $p_i$ .	Price
4	FM	The market operator allocates $C$ price-wise among the $S_i$ .	–
5	RTM	All energy is transacted at market price the next day.	–

**4.3.3.3 Bertrand-Edgeworth competition** In the previous section, we described two market formats in which price bidding is preceded by capacity competition. We now apply Bertrand-Edgeworth competition to assemble the tools needed to analyze **M1** and **M2** in terms of market efficiency and system reliability.

Capacity followed by price competition has been the subject of a large body of literature, beginning with [51] where it was shown that the duopoly case yields Cournot outcomes. For analytical tractability, we also focus primarily on duopolies. We will give new theoretical results for case in which demand is random, which is arguably the only realistic way of modeling demand for electric power in the RTM.

**4.3.3.3.1 Allocation and payoff rules** We now detail price competition for a generic quantity  $B$  under the assumption that firm capacities  $S_i$  have already been set. Demand at each time step is distributed according the linear program

$$\begin{aligned}
& \min_{X_i} && \sum_{i=1}^{N+1} p_i X_i \\
& \text{s. t.} && \sum_{i=1}^{N+1} X_i = B \\
& && 0 \leq X_i \leq S_i
\end{aligned}$$

Table 4: Sequence of events in **M2**

Stage	Time	Description	Competition
1	FM	The system operator purchases $\Delta - C$ generator reserves.	–
2	RTM	Each firm $i$ specifies a maximum capacity $S_i \leq \bar{S}_i$ to the system operator.	Capacity
3	RTM	Each firm $i$ offers an energy imbalance fee $p_i$ .	Price
4	RTM	The market operator allocates a random amount of energy price-wise among the $S_i$ .	–
5	RTM	Energy from NGR $i$ is purchased at $p_i + m$ plus the market price, and any additional reserves required are purchased from generators.	–

Table 5: Comparison of **M1** and **M2**

Format	Demand	Type	Price strategy	Demand
<b>M1</b>	Capacity	Deterministic	Capacity premium	FM
<b>M2</b>	Energy	Random	Imbalance fee	RTM

$X_i$  is the amount given to firm  $i$ . Firm  $N + 1$  corresponds to generation reserves  $p_{N+1} = R$  and  $S_{N+1} = \infty$ , which implies that  $p_i \leq R$  if firm  $i$  is to receive any payment for reserves; for this reason,  $R$  is sometimes referred to as the reservation utility. Although it is an important practical matter to specify allocation mechanisms for ties, it does not factor into our analysis and so is omitted.

The payment to firms  $i$  is then  $p_i X_i$ ; note that the vector  $X$  may not be unique, as multiple vectors  $X$  may solve the above optimization when  $p_i = p_j$  for some  $i, j$ . Define the set

$$Y_i(p) = \{j \mid p_j < p_i\}, \quad (13)$$

and let  $\mu_i$  be firm  $i$ 's mixed strategy. We may write firm  $i$ 's profits

$$\begin{aligned} \rho_i(\mu; B) &= \int_p \left( \prod_j \mu_j(p_j) \right) p_i \\ &\quad \times \min \left\{ S_i, \left( B - \sum_{j \in Y_i(p)} S_j \right)^+ \right\} dp. \end{aligned} \quad (14)$$

We will slightly abuse notation in denoting the payoff of a pure strategy  $\rho(p_i, \mu_{-i}; B)$ . We will make the dependence on  $R$  and  $S$  explicit when denoting payoffs at equilibrium:

$$\rho_i^*(B, R, S) = \rho_i(\mu; B). \quad (15)$$



Later, we will identify  $B$  as both a deterministic capacity in **M1** and a random energy in **M2**, in which cases prices respectively correspond to capacity premiums and imbalance fees. In the latter case, we assume that  $B$  is drawn from the probability distribution  $f$ , which we approximate as having the support  $[0, \infty)$ . When we discuss capacity competition in Section 4.3.3.2, we will assume the market operator wishes to procure  $\Delta$  total reserves, and that  $C$  is the quantity they “leave” to the NGRs. We will write  $f_C$  to denote  $f$  with its domain shifted by  $\Delta - C$  as a result of absorption by conventional reserves, i.e.

$$f_C(B) = f(B + \Delta - C) + \delta_0(B) \int_0^{\Delta-C} f(x)dx, \quad (16)$$

where  $B \geq 0$  and  $\delta_0$  is the Dirac delta function centered at zero. The corresponding cumulative distribution is given by  $F_C(B) = F(B + \Delta - C)$ , again for  $B \geq 0$ .

**Price competition in M1** When capacity is procured in the FM as in **M1**,  $B$  is deterministic and firm payoffs resulting from price competition are summarized by the following proposition from [12]:

**Proposition 1** [12]

1. If  $B \leq -S_j + \sum_i S_i$  for all  $j$ , the pure-strategy equilibrium is  $p_i = 0$ ,  $i = 1, \dots, N$ , and each firms payoff is zero.
2. If  $B \geq \sum_i S_i$ , the pure strategy Nash equilibrium is  $p_i = R$ ,  $i = 1, \dots, N$ , and each firms payoff is  $RS_i$ .
3. If  $-S_j + \sum_i S_i < B < \sum_i S_i$  for some  $j$ , a mixed strategy exists, and the payoff of firm  $j$  is

$$\rho_j^*(B, R, S) = \frac{R \min\{S_j, B\}}{\min\{S_{\max}, B\}} \left( B + S_{\max} - \sum_i S_i \right)$$

When there are two firms in this case and  $S_1 \geq S_2$ , the mixed equilibrium strategies are given by

$$\begin{aligned} \mu_1(x) &= \frac{R(B - S_2)S_2}{\min\{B, S_1\}(\min\{B, S_1\} + S_2 - B)x^2}, \\ &\quad x \in [L, R) \\ \mu_2(x) &= \frac{R(B - S_2)}{(\min\{B, S_1\} + S_2 - B)x^2}, \quad x \in [L, R]. \end{aligned}$$

In the two firm case, we denote the larger and smaller firms' equilibrium payoffs by  $\bar{\rho}(B, R, S)$  and  $\underline{\rho}(B, R, S)$ , respectively.

**Price competition in M2** We now assume  $B$  represents energy realized in the RTM, and is thus random. Because there is no danger of overlapping notation, we will also denote mixed strategies under random demand by  $\mu$ . Each firm seeks to maximize its expected payoffs, which by Fubini's theorem satisfies

$$\pi_i(\mu) = \mathbb{E}[\rho_i(\mu; B)].$$

We remark that this definition is essentially the Bayesian equilibrium when  $B$  is the only unknown parameter and all players share the same knowledge [40]. We first address existence of equilibria.

**Proposition 2** *No pure-strategy Nash equilibrium exists.*

**Proposition 3** *A mixed-strategy equilibrium exists.*

Let  $U_i$  and  $L_i$  be the upper and lower supports of  $\mu_i$ . By the definition of mixed-strategy equilibrium [60], there exists a  $\pi_i^*$  and a subset  $P_i \subseteq [L_i, U_i]$ ,  $\mu_i(P_i) = 1$  for which

$$\begin{aligned}\pi_i(p_i, \mu_{-i}) &\leq \pi_i^*(C, R, S) & \forall p_i \in [U_i, L_i], \\ \pi_i(p_i, \mu_{-i}) &= \pi_i^*(C, R, S) & \forall p_i \in P_i.\end{aligned}$$

Because  $\rho_i(p_i, \mu_{-i}; B)$  is continuous at  $p_i$  if  $\mu_{-i}$  has no atom there, so is  $\pi_i(p_i, \mu_{-i})$ , and thus  $\pi_i(U_i, \mu_{-i}) = \pi_i^*$  if  $\mu_{-i}$  has no atom at  $U_i$ .

**Lemma 1** *Behavior under random demand*

1. *For any  $[L_j, U_j]$ , there are at least two firms  $i$  for which  $\mu_i(p) > 0$  for all  $p \in [L_j, U_j]$ .*
2. *Only one atom can exist, and if one does, it must be at  $U$ .*
3.  $U = R$ .

Note that if firm  $i$  has maximal capacity, its expected equilibrium profits are identical if  $B$  is announced before or after price competition:

$$\begin{aligned}\pi_i^*(C, R, S) &= E[\rho_i^*((B - \Delta + C)^+, R, S)] \\ &= R \left[ \int_0^{S_i} B f_C \left( B + \sum_{j \neq i} S_j \right) dB \right. \\ &\quad \left. + S_i \int_{S_i}^{\infty} f_C \left( B + \sum_{j \neq i} S_j \right) dB \right]\end{aligned}$$

We henceforth restrict our attention the more tractable two firm case. Define

$$\begin{aligned}\bar{S}_i &= \int_0^{S_i} B f_C(B) dB + S_i \int_{S_i}^{\infty} f_C(B) dB \\ \phi &= \int_0^{S_2} B f_C(B + S_1) dB + S_2 \int_{S_2}^{\infty} f_C(B + S_1) dB.\end{aligned}$$

For notational simplicity, we now respectively denote the optimal payoff of the larger and smaller firm  $\bar{\pi}(C, R, S)$  and  $\underline{\pi}(C, R, S)$ , and in this section assume  $S_1 \geq S_2$ .

**Lemma 2** *Only the larger firm may have an atom at  $R$ .*

Since only firm one may have an atom at  $R$ , its profits are given by

$$\begin{aligned}\bar{\pi}(C, R, S) &= R \left[ \int_0^{S_1} B f_C(B + S_2) dB \right. \\ &\quad \left. + S_1 \int_{S_1}^{\infty} f_C(B + S_2) dB \right],\end{aligned}$$

and, using the above reasoning, those of two by

$$\begin{aligned}\underline{\pi}(C, R, S) &= \bar{\pi}(C, R, S) \frac{\pi_2(L, \mu_1)}{\pi_1(L, \mu_2)} \\ &= \bar{\pi}(C, R, S) \frac{\bar{S}_2}{\bar{S}_1}\end{aligned}$$

As in the previous section, we may now explicitly characterize the mixed-equilibrium strategies  $\mu_1$  and  $\mu_2$  in the same fashion as before. Let  $\Upsilon_i$  be the cumulative distribution of the mixed-strategy of firm  $i$ .

**Proposition 4**

$$\mu_1(x) = \frac{\underline{\pi}(C, R, S)}{(\bar{S}_2 - \phi) x^2} x \in [L, R) \quad (17)$$

$$\mu_2(x) = \frac{R\bar{\pi}(C, R, S)}{(R\bar{S}_1 - \bar{\pi}(C, R, S)) x^2} x \in [L, R] \quad (18)$$

We write  $\mu_1(p; R, S, C)$  and  $\mu_2(p; R, S, C)$  when we wish to make the dependency on  $R, S$  and  $C$  explicit.

**4.3.3.3.2 Capacity competition** We now turn to the determination of NGR capacities,  $S$ , in **M1** and **M2**. Suppose the system operator's total demand for reserves is  $\Delta$ . The cost of generator reserve capacity is given by the monotonically increasing function  $r(x)$ , for which we assume that  $r(0) = 0$  and  $0 < \lim_{x \rightarrow 0} r(x)/x < \infty$ . Define  $\gamma_i$  to be firm  $i$ 's capacity opportunity cost.

We model participation of the system operator as a Stackelberg leader [40], in which first the portion of  $\Delta$  to be left to NGRs,  $C$ , is announced, after which the NGRs engage in capacity and then price competition. The Stackelberg leader concept has been applied in other power system settings, such as capacity expansion [74] and competition among generators when one is substantially larger than the others [79].

We assume that the market operator will only patronize NGRs for energy after it has exhausted the  $\Delta - C$  generator reserves it has already procured. While this may neglect the superior performance of NGRs, it is economically inefficient for the system operator to not utilize resources it has already paid for. This assumption has little effect in **M1**, in which firms are paid for capacity, but plays a more substantial role in the energy payments of **M2**.

**M1** In **M1**,  $C$  is directly requested of the NGRs, and so is equal to the total demand. Define the price of generator reserve capacity to be  $R_1(x) = r(x)/x$ . The net cost to the system operator is then

$$\begin{aligned}\Pi(C, S) &= r(\Delta - C) + \bar{p}(C, R_1(\Delta - C), S) \\ &\quad + \underline{p}(C, R_1(\Delta - C), S),\end{aligned}$$

and the net profits of storage  $i$  are given by

$$\begin{aligned}\tilde{\rho}_i(C, S) &= \begin{cases} \bar{p}(C, R_1(\Delta - C), S_i, S_{-i}) & S_i \geq S_{-i} \\ \underline{p}(C, R_1(\Delta - C), S_{-i}, S_i) & S_i < S_{-i} \end{cases} \\ &\quad - \gamma_i S_i\end{aligned}$$

These utilities define a game in which the market operator optimizes the requested capacity  $C$ , and subsequently the NGRs engage in a capacity subgame.

In the absence of physical NGR capacities  $\mathbb{S}$ , the pure strategy capacity equilibrium is given by [12]:

$$\sum_i S_i = C \quad \text{and} \quad \frac{R_1(\Delta - C) - \gamma_i}{2R_1(\Delta - C) - \gamma_i} C \leq S_i \leq S_{-i} \quad (19)$$

for some  $i$ . With physical capacities, the equilibrium is characterized following proposition:

**Proposition 5** *The equilibria are given by the following cases:*

- If  $\mathbb{S}_1 + \mathbb{S}_2 \leq C$ , then  $S_i = \mathbb{S}_i$
- If  $\mathbb{S}_1 + \mathbb{S}_2 > C$  and solutions exist satisfying (19) and  $S \leq \mathbb{S}$ , then all such solutions are equilibria.
- Suppose  $\mathbb{S}_1 + \mathbb{S}_2 > C$  but no solutions exist satisfying both (19) and  $S \leq \mathbb{S}$ . Without loss of generality, additionally suppose that  $\mathbb{S}_1 < \frac{R_1(\Delta - C) - \gamma_1}{2R_1(\Delta - C) - \gamma_1} C$ . Then  $S_1 = \mathbb{S}_1$  and  $S_2 = C - \mathbb{S}_1$  is the equilibrium.

Because this is a capacity regime in which a pure strategy price equilibrium exists, equilibrium profits always satisfy

$$\tilde{\rho}_{\tau(i)}(C, S) = (R_1(\Delta - C) - \gamma_i) S_i. \quad (20)$$

We must ensure that the market operator does not choose  $C$  greater than that which can be provided, both due to absolute capacity constraints and market participation; if  $R_1(\Delta - C) < \gamma_i$ , firm  $i$  offers zero capacity to the market, which is modeled through the indicator function

$$\delta_i(C) = \begin{cases} 1 & R_1(\Delta - C) \geq \gamma_i \\ 0 & \text{otherwise} \end{cases}. \quad (21)$$

The market operator then solves

$$\max_{C, S, k} \quad r(\Delta - C) + R_1(\Delta - C)C \quad (22)$$

$$\text{s.t.} \quad 0 \leq C \leq \Delta \quad (23)$$

$$C = \sum_i S_i \quad (24)$$

$$0 \leq S_i \leq \delta_i(C) \mathbb{S}_i \quad (25)$$

$$\min \left\{ \frac{R_1(\Delta - C) - \gamma_k}{2R_1(\Delta - C) - \gamma_k} C, \mathbb{S}_k \right\} \leq S_k \quad (26)$$

$$S_k \leq S_{-k} \quad (27)$$

$$k \in \{1, 2\} \quad (28)$$

In this optimization, the system operator essentially selects which firm has maximal capacity via the integer variable  $k$ . The constraints (26) and (27) ensure that the resulting solution is an equilibrium satisfying one of the cases of Prop. 5. If multiple optimal capacities exist, one is chosen at random. The resulting optimal vector of capacities as a function of  $\mathbb{S}$  by  $\Phi^1(\mathbb{S})$ .

When  $r$  is a linear function, any feasible  $C$  is an equilibrium. When all players act simultaneously, this is not true, but  $C = 0$  and  $S_i = 0$  remains a trivial equilibrium; intuitively, this is because the system operator will attempt to lower storage prices by reducing its demand, while at the same time the NGRs

attempt to increase their prices by making their own capacities scarce. The same is true in both the Stackelberg and simultaneous cases if  $r$  increases sublinearly. Contrarily, if  $r$  increases superlinearly, the Stackelberg leader requests the maximum available or useful capacity from NGRs. It is most realistic to assume, as in primary energy markets, that the cost of generator reserves is increasing and strictly convex (e.g. convex quadratic [13]), indicating a high utilization of NGRs.

**M2** We now consider capacity commitment when energy is procured from the NGRs in the RTM through competitive pricing. We must first define a price for energy from generator reserves. Let  $R_2(\Delta - C)$  be the expected effective cost of energy from generator reserves. Since the energy imbalance realization cannot be known until after generators commit capacity, and to make  $R_2(\Delta - C)$  bounded, it is defined as

$$R_2(x) = \frac{r(x)}{E[\min\{B, x\}]} \quad (29)$$

The market operator's costs are

$$\begin{aligned} \Pi(C, S) = & r(\Delta - C) + \bar{\pi}(C, R_2(\Delta - C), S) \\ & + \underline{\pi}(C, R_2(\Delta - C), S), \end{aligned}$$

and the net profits of storage  $i$  are given by

$$\begin{aligned} \tilde{\pi}_i(C, S) = & \begin{cases} \bar{\pi}(C, R_2(\Delta - C), S_i, S_{-i}) & S_i \geq S_{-i} \\ \underline{\pi}(C, R_2(\Delta - C), S_{-i}, S_i) & S_i < S_{-i} \end{cases} \\ & - \gamma_i S_i. \end{aligned}$$

Note that  $C$  has a slightly different meaning here than in **M1**: it is now the leftover capacity not procured via traditional reserves, rather than that requested from NGRs. We may thus interpret  $f_C$  as the distribution of the energy imbalance minus the amount absorbed by generator reserves,  $\Delta - C$ . We first establish properties of the NGR payoff functions.

**Lemma 3**  $\bar{\pi}(C, R_2(\Delta - C), S) - \gamma_i S_i$  and  $\underline{\pi}(C, R_2(\Delta - C), S) - \gamma_{-i} S_{-i}$  are respectively strictly concave in  $S_i$  and strictly quasiconcave in  $S_{-i}$ .

We are now prepared to discuss the equilibria themselves. Because  $\pi_1$  and  $\pi_2$  are quasiconcave and differentiable, Nash equilibria may occur where their derivatives are zero or at their capacities,  $\mathbb{S}$ . Let

$$\begin{aligned} \Lambda_i^1 &= (F_C^{-1}(1 - \gamma_i/R_2(\Delta - C)))^+ \\ &= (C - \Delta + F^{-1}(1 - \gamma_i/R_2(\Delta - C)))^+. \end{aligned} \quad (30)$$

Setting the derivative of  $\bar{\pi}(C, R_2(\Delta - C), S) - \gamma_i S_i$  with respect to  $S_i$  equal to zero, we have that at equilibrium  $S_i = \min\{\Lambda_i^1 - S_{-i}, S_i\}$  if  $S_i \geq S_{-i}$ .

When  $S_i < S_{-i}$ , the first derivative is given by

$$\begin{aligned} \frac{d\pi}{dS_2}(C, R_2(\Delta - C), S) = & \frac{1}{\bar{S}_1} [(1 - F_C(S_2)) \\ & \times \bar{\pi}(C, R_2(\Delta - C), S) \\ & + R_2(\Delta - C)(F_C(S_2) \\ & - F_C(S_1 + S_2)) \bar{S}_2]. \end{aligned} \quad (31)$$

Now substitute  $\min \{\Lambda_i^1 - S_{-i}, \mathbb{S}_i\}$  for  $S_i$  in (31), and let  $\Lambda_{-i}^2$  be the value of  $S_{-i}$  satisfying the nonlinear equation

$$\frac{d\pi}{dS_2}(C, R_2(\Delta - C), \min \{\Lambda_i^1 - S_{-i}, \mathbb{S}_i\}, S_{-i}) = \gamma_{-i}.$$

$S_{-i}$  at equilibrium is then  $\min \{\Lambda_{-i}^2, \mathbb{S}_{-i}\}$ . We denote the pair

$(\min \{\Lambda_i^1 - \min \{\Lambda_{-i}^2, \mathbb{S}_{-i}\}, \mathbb{S}_i\}, \min \{\Lambda_{-i}^2, \mathbb{S}_{-i}\})$  by  $\hat{S}^i(C)$ . We will use the following definition to analyze uniqueness:

$$\begin{aligned} \tilde{S}_{-i}(S_i) = & \min \left\{ \operatorname{argmax}_{S_{-i}} \pi(C, R_2(\Delta - C), S_i, S_{-i}) \right. \\ & \left. - \gamma_{-i} S_{-i}, \mathbb{S}_{-i} \right\}. \end{aligned}$$

$\tilde{S}_{-i}(S_i)$  is well-defined due to the quasiconcavity of  $\pi(C, R_2(\Delta - C), S_i, S_{-i}) - \gamma_{-i} S_{-i}$  with respect to  $S_{-i}$ .

**Lemma 4**  $\hat{S}^i(C)$  is unique if  $\tilde{S}_{-i}(\Lambda_i^1/2) < \Lambda_i^1/2$ .  $\gamma_i < \gamma_{-i}$  is a sufficient condition for this.

Lemma 4 only tells us when  $\hat{S}^i(C)$  are unique, but not that they are actual equilibria. To be equilibria, they must satisfy the definition of a Nash equilibrium:

$$\hat{S}_i^i(C) = \operatorname{argmax}_{S_i} \tilde{\pi}_i^*(C, S_i, S_{-i}) \quad (32)$$

$$\hat{S}_{-i}^i(C) = \operatorname{argmax}_{S_{-i}} \tilde{\pi}_{-i}^*(C, S_{-i}, S_i) \quad (33)$$

Composing the above developments, we have the following result.

**Theorem 1** *The  $\hat{S}^i(C)$  are the only possible pure-strategy equilibria.*

Since  $\tilde{\pi}_i$  is continuous, a mixed-strategy equilibrium must exist [41] even if each  $\hat{S}^i(C)$  fails to be an equilibrium. We comment that if  $\gamma_i > R_2(\Delta - C)$  for both  $i$ ,  $S = 0$  is the only equilibrium, and if  $\gamma_i \leq R_2(\Delta - C)$  and  $\gamma_{-i} > R_2(\Delta - C)$ ,  $S_i = \min \{\Lambda_i^1, \mathbb{S}_i\}$  and  $S_{-i} = 0$  is the only equilibrium.

We now discuss the resulting Stackelberg game in **M2**. We must have  $C \leq S_1 + S_2$  to satisfy the minimum reserve requirement as well as  $0 \leq S_i \leq \mathbb{S}_i$ . The system operator thus solves

$$\min_C \quad r(\Delta - C) + \Theta \quad (34)$$

$$\text{s.t.} \quad 0 \leq C \leq \Delta \quad (35)$$

$$C \leq \min \left\{ \sum_i \hat{S}_i^1(C), \sum_i \hat{S}_i^2(C) \right\} \quad (36)$$

$$\Phi = \begin{cases} \pi(C, R_2(\Delta - C), \hat{S}^i(C)) \\ \quad + \pi(C, R_2(\Delta - C), \hat{S}^i(C)) \\ \text{if } i \text{ exists} \\ \frac{1}{2} \sum_{i=1,2} \pi(C, R_2(\Delta - C), \hat{S}^i(C)) \\ \quad + \pi(C, R_2(\Delta - C), \hat{S}^i(C)) \\ \text{if both exist} \\ \infty \text{ if none exist} \end{cases} \quad (37)$$

The final constraint is equal to the payment to the NGRs if only one equilibrium exists, is the average over the two equilibria if both exist, and is infinite if none exist; we remark that this constraint is overly stringent, because a mixed-strategy capacity equilibrium exists in the latter scenario, which would not be infinitely costly. Despite the fact that this optimization problem does not have a particularly nice analytical structure, it only has one variable, and is thus relatively easy to solve. We denote the optimal vector of capacities as a function of  $\mathbb{S}$  by  $\Phi^2(\mathbb{S})$ .

**4.3.3.4 Analysis of energy markets** The preceding results are clearly too stylized to provide quantitative insights, but do enable us to make a number of qualitative predictions regarding various market design choices.

**4.3.3.4.1 Analytical observations** According to Prop. 1 of [12], firms will set prices consistently in **M1** (competition on the basis of capacity premiums), but always at the capacity reservation price,  $R_1(\Delta - C)$ , the maximally inefficient outcome from the system operator's perspective. Prop. 2 says that in **M2** (competition on the basis of energy imbalance fees), firms will price less consistently because they will always be in a mixed-strategy regime; however, they will also set prices lower than the energy reservation utility,  $R_2(\Delta - C)$ , which is more desirable to the system operator.

In **M1**, one can expect less consistent capacity commitment: although the total amount offered will match demand as closely as possible, fairness among firms can vary widely due to the potentially infinite range of equilibria, cf. Prop. 5 and [12]. The situation is more complicated in **M2**. Theorem 1 says that under certain conditions, at most two pure-strategy equilibria can exist, but in some scenarios there may only be mixed equilibria.

The fact that a finite number of pure strategy capacity equilibria exist in **M2**, and implies that firms should dependably commit capacity when they exist. It is primarily a consequence of the fact that analytic functions have isolated zeros, but is reminiscent of the main result of [50], in which it is also shown that uncertainty reduces the number of pure-strategy equilibria. Unfortunately, this is only a beneficial characteristic when pure-strategy equilibria exist. While mixed strategy behaviors are tolerable in the pricing stage because they do not affect the actual resource available, we regard them as undesirable in the capacity stage, in which dependable behavior is more important. In this regard, **M2** offers lower volatility when pure-strategy equilibria exist, but may be unacceptable when they fail to exist.

**4.3.3.4.2 Single stage simulation** We now examine the capacity equilibria over influential parameter ranges. We consider scaled instances with  $\Delta = C = 1$ , where random, positive energy follows the distribution  $f(B) = \sqrt{2/\pi}e^{-B^2/2}$ . Suppose further that procuring  $x$  generator reserve capacity costs  $r(x) = ax^2 + x$ . The price of reserve capacity is then  $R_1(x) = r(x)/x$  and the price of expected reserve energy  $R_2(x) = r(x)/E[\min\{B, x\}]$ . Since  $\Delta = C$ ,  $R_1 = 1$  and

$$\begin{aligned} R_2 &= \lim_{x \rightarrow 0} \frac{r(x)}{E[\min\{B, x\}]} \\ &= \lim_{x \rightarrow 0} \frac{ax + 1}{1 - F(x)} \text{ by l'Hôpital's rule} \\ &= 1. \end{aligned}$$

We are interested in equilibrium existence in **M2**, and profitability in both formats. In all cases, each  $\hat{S}_i(C)$  satisfies the condition of Lemma 4, and is thus unique. In each of the first two figures, the outcome

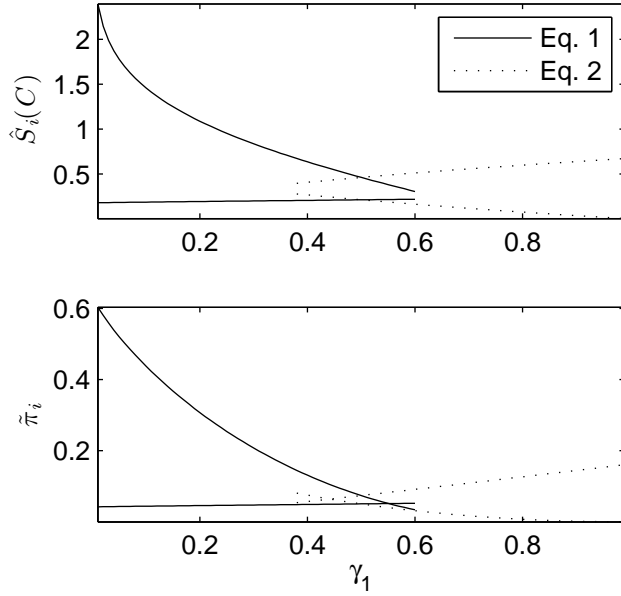


Figure 11: Each NGR's equilibrium capacity and payoffs for  $\gamma_2 = 0.5$  and  $\mathbb{S}_1 = \mathbb{S}_2 = \infty$  in **M2**.

for both NGRs in each equilibrium is shown.

The upper plot of Fig. 11 shows the **M2** capacity equilibrium as firm one's opportunity cost is varied between zero and one while the other's remains fixed at 0.5. If one firm has a substantially smaller opportunity cost than the other, only one equilibrium will exist, in which the smaller opportunity cost firm commits more capacity. If the opportunity costs are relatively close, two equilibria exist corresponding to both capacity orderings.

Fig. 12 shows the resulting equilibria as  $\mathbb{S}_1$  is varied while  $\mathbb{S}_2 = \infty$ ,  $\gamma_1 = 0.4$ , and  $\gamma_2 = 0.6$ . When  $\mathbb{S}_1$  is large, only one equilibrium exists in which firm one commits more capacity. As  $\mathbb{S}_1$  is decreased, this equilibrium eventually ceases to exist, and another with opposite capacity ordering appears.

These simulations show that when firms are nearly symmetric, either can play the role of the larger firm, and even if they are not symmetric, hard physical capacity constraints can force the same indifference by prohibiting one firm from being the larger of the two.

In Fig. 13, we observe the total NGR profits as functions of  $\alpha$  when opportunity costs are  $\gamma_1 = 0.3\alpha$  and  $\gamma_2 = 0.7\alpha$ . Because firms charge the reservation utility  $R_1$  in **M1** but play mixed pricing strategies in **M2**, NGR profits are substantially higher in **M1**.

**4.3.3.5 Conclusion and future work** NGRs such as energy storage and aggregations of shiftable loads could soon compete alongside traditional power plants in reserve and regulation markets. Using capacity followed by price competition, we have compared competition between NGRs when they are paid for capacity versus mileage. In the case of the former, pure capacity commitment strategies always exist, but there may be an infinite number of them. In the case of the latter, at most two pure strategy equilibria exist under mild conditions, but in some instances there may be none. Through analysis and numerical simulation, we observe that NGRs profit substantially more under capacity payments because they



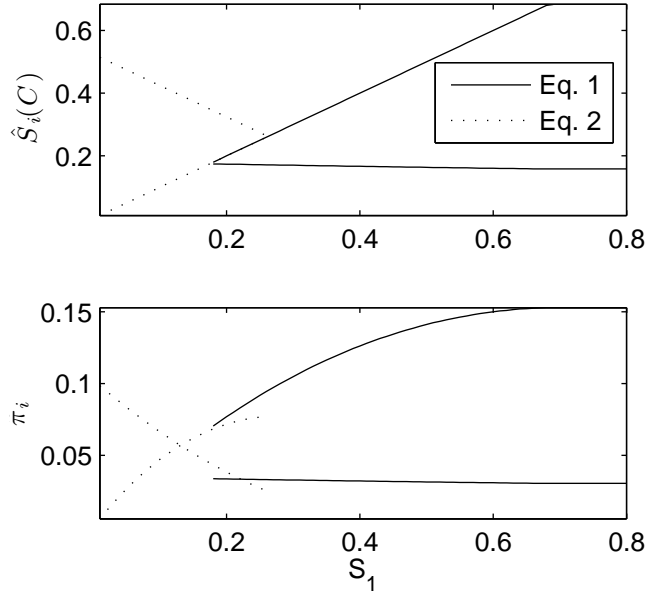


Figure 12: Each NGR's equilibrium capacity and payoffs for  $\gamma_1 = 0.4$ ,  $\gamma_2 = 0.6$ , and  $S_2 = \infty$  in **M2**.

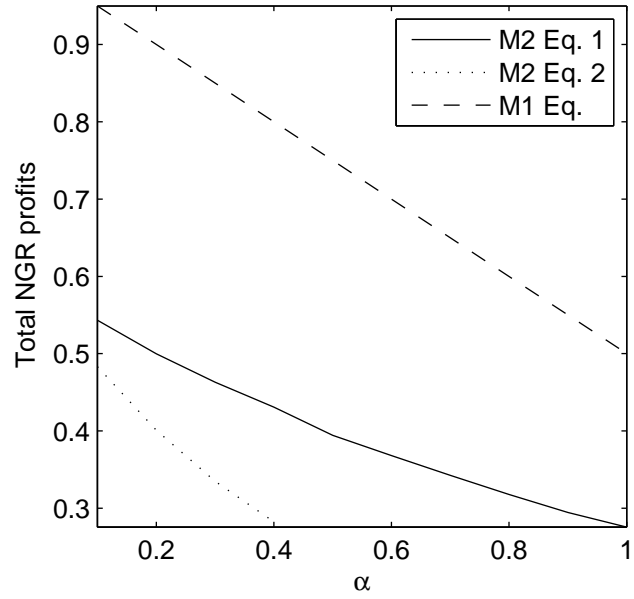


Figure 13: Total NGR profits in **M1** and for each equilibrium in **M2** as a function of  $\alpha$ , where  $\gamma_1 = 0.3\alpha$  and  $\gamma_2 = 0.7\alpha$ .

consistently charge the reservation utility, whereas in **M2** mixed pricing strategies are played, resulting in reduced profits.

We interpret these results as follows. When firms profit from energy imbalances, a fundamentally random quantity, preparing for all possible outcomes specifies a small number of strategies, while when they profit from a known capacity demand, a range of behaviors are sufficient; this result resembles that of [50] for supply function equilibria. In the context of zero-mean energy markets, it suggests that incorporating random energy payments may yield more predictable behaviors from NGR participants. Energy payments also induce more intense competition at the pricing stage because NGRs cannot tune their capacities to a random demand. This leads to lower NGR profits and hence lower prices for end consumers. Ultimately, the recommendation of this work is that hybrid capacity and mileage payments be used, which is a current direction of system operators [25].

The heavily stylized nature of our analysis leaves many venues for future work. In particular, comparison of uniform versus discriminatory payments mechanisms, incorporation of physical losses, and  $n$ -firm analysis with barriers to entry are potential topics.

## References

- [1] *Distribution Data Book*. Number GET-1008L. General Electric, 1972.
- [2] Ieee guide for loading mineral-oil-immersed transformers, 1996.
- [3] *Distribution Transformers*. ABB, Business Area Distribution Transformers, P.O. Box 8131, CH - 8050 Zurich, Switzerland, 2001.
- [4] Nov 2012.
- [5] 2014 general rate case phase I prepared testimony exhibit PG&E-4: Electric distribution, November 2012. (workpaper pages can be requested separately from PG&E).
- [6] Measurement and instrumentation data center, Nov 2012.
- [7] Revised California energy demand forecast 2012-2022 volume 2: Electricity demand by planning area. Staff Rep. CEC-200-2012-001-SD-V2, California Energy Commission, February 2012.
- [8] California energy demand 2014-2024 preliminary forecast volume 1: Statewide electricity demand, end-user natural gas demand, and energy efficiency. Staff Rep. CEC-200-2013-004-SD-V1, California Energy Commission, May 2013.
- [9] California ISO open access same-time information system (OASIS), April 2013.
- [10] California ISO peak load history 1998 through 2012, July 2013.
- [11] Quality controlled local climatological data, July 2013.
- [12] Daron Acemoglu, Kostas Bimpikis, and Asuman Ozdaglar. Price and capacity competition. *Games and Economic Behavior*, 66(1):1 – 26, 2009.
- [13] E.H. Allen and M.D. Ilic. Reserve markets for power systems reliability. *Power Systems, IEEE Transactions on*, 15(1):228 –233, Feb. 2000.
- [14] C. Alvarez, A. A. Gabaldón, and A. Molina. Assessment and simulation of the responsive demand potential in end-user facilities: Application to a university customer. *IEEE Transactions on Power Systems*, 19(2):1223–1231, 2004.
- [15] María Ángeles de Frutos and Natalia Fabra. Endogenous capacities and price competition: The role of demand uncertainty. *International Journal of Industrial Organization*, 29(4):399 – 411, 2011.
- [16] Xiaoqing Bai, Hua Wei, Katsuki Fujisawa, and Yong Wang. Semidefinite programming for optimal power flow problems. *International Journal of Electrical Power and Energy Systems*, 30(6-7):383 – 392, 2008.
- [17] M.L. Baughman, S.N. Siddiqi, and J.W. Zarnikau. Advanced pricing in electrical systems. I. Theory. *Power Systems, IEEE Transactions on*, 12(1):489 –495, Feb. 1997.
- [18] A.W. Berger and F.C. Schweppe. Real time pricing to assist in load frequency control. *Power Systems, IEEE Transactions on*, 4(3):920–926, Aug. 1989.

- [19] Roy Billinton and Ronald N. Allan. *Reliability evaluation of power systems*. Springer, second edition, 1996.
- [20] F. Borrelli, A. Bemporad, and M. Morari. *Predictive control for linear and hybrid systems*. Cambridge University Press, 2011.
- [21] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*, volume 15 of *Studies in Applied Mathematics*. SIAM, Philadelphia, PA, 1994.
- [22] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, New York, NY, USA, 2004.
- [23] Richard J Bravo. Dynamic performance of residential loads. Master’s thesis, California State University, Long Beach, 2012.
- [24] CAISO. Flexible ramping products: Revised draft final proposal. Technical report, California Independent System Operator, Aug. 2012.
- [25] CAISO. Pay for performance regulation: Draft final proposal. Technical report, California Independent System Operator, Mar. 2012.
- [26] California ISO. Regulation energy management draft final proposal, Jan. 2011.
- [27] D.S. Callaway and I. A. Hiskens. Achieving controllability of electric loads. *Proceedings of the IEEE*, 99(1):184–199, January 2011.
- [28] J.M. Carrasco, L.G. Franquelo, J.T. Bialasiewicz, E. Galvan, R.C.P. Guisado, Ma.A.M. Prats, J.I. Leon, and N. Moreno-Alfonso. Power-electronic systems for the grid integration of renewable energy sources: A survey. *Industrial Electronics, IEEE Transactions on*, 53(4):1002–1016, Jun. 2006.
- [29] John Carruthers. personal communication, 2013.
- [30] Michael A Cohen, May 2013.
- [31] First Hydro Company. Dinorwig power station, 2009.
- [32] Beacon Power Corp. Energy storage: Regulation issues. ERCOT Emerging Technologies Working Group, 2011.
- [33] F.Y. Edgeworth. The pure theory of monopoly. In *Papers relating to political economy*, volume 1, pages 111–142. Macmillan and Co., Ltd., 1925.
- [34] O.I. Elgerd and C.E. Fosha. Optimum megawatt-frequency control of multiarea electric energy systems. *Power Apparatus and Systems, IEEE Transactions on*, PAS-89(4):556–563, Apr. 1970.
- [35] J.H. Eto, J. Nelson-Hoffman, E. Parker, C. Bernier, P. Young, D. Sheehan, J. Kueck, and B. Kirby. Demand response spinning reserve demonstration: Phase 2 findings from the summer of 2008. Technical Report LBNL-2490E, Lawrence Berkeley National Laboratory, 2009.
- [36] Natalia Fabra, Nils-Henrik von der Fehr, and David Harbord. Designing electricity auctions. *The RAND Journal of Economics*, 37(1):23–46, 2006.

- [37] Federal Energy Regulatory Commission. Order no. 755: Frequency regulation compensation in the organized wholesale power markets, Oct. 2011.
- [38] Luis Fernandes. personal communication, 2012.
- [39] C.E. Fosha and O.I. Elgerd. The megawatt-frequency control problem: A new approach via optimal control theory. *Power Apparatus and Systems, IEEE Transactions on*, PAS-89(4):563–577, Apr. 1970.
- [40] Drew Fudenberg and Jean Tirole. *Game Theory*. MIT Press, 1991.
- [41] I. L. Glicksberg. A further generalization of the Kakutani fixed point theorem, with application to nash equilibrium points. *Proceedings of the American Mathematical Society*, 3(1):170–174, 1952.
- [42] H.B. Gooi, D.P. Mendes, K.R.W. Bell, and D.S. Kirschen. Optimal scheduling of spinning reserve. *Power Systems, IEEE Transactions on*, 14(4):1485–1492, Nov. 1999.
- [43] Richard J. Green and David M. Newbery. Competition in the British electricity spot market. *Journal of Political Economy*, 100(5):pp. 929–953, 1992.
- [44] D.A. Halamay, T.K.A. Brekken, A. Simmons, and S. McArthur. Reserve requirement impacts of large-scale integration of wind, solar, and ocean wave power generation. *Sustainable Energy, IEEE Transactions on*, 2(3):321–328, Jul. 2011.
- [45] Anderson Hoke, Rebecca Butler, Joshua Hambrick, and Benjamin Kroposki. Steady-state analysis of maximum photovoltaic penetration levels on typical distribution feeders. *Sustainable Energy, IEEE Transactions on*, 4(2):350–357, 2013.
- [46] M. Ilic, P. Skantze, C.N. Yu, L. Fink, and J. Cardell. Power exchange for frequency control (PXFC). In *Power Engineering Society 1999 Winter Meeting, IEEE*, volume 2, pages 809 – 819, Jan. 1999.
- [47] R.A. Jabr. Radial distribution load flow using conic programming. *Power Systems, IEEE Transactions on*, 21(3):1458–1459, Aug. 2006.
- [48] J.R. Green and J.J. Laffont. *Incentives in Public Decision Making*. North-Holland, Amsterdam, 1979.
- [49] B. Kirby. Spinning reserve from responsive loads. Technical Report ORNL/TM-2003/19, Oak Ridge National Laboratory, 2003.
- [50] Paul D. Klemperer and Margaret A. Meyer. Supply function equilibria in oligopoly under uncertainty. *Econometrica*, 57(6):1243–1277, 1989.
- [51] David M. Kreps and Jose A. Scheinkman. Quantity precommitment and Bertrand competition yield Cournot outcomes. *The Bell Journal of Economics*, 14(2):326–337, 1983.
- [52] Prabha Kundur. *Power system stability and control*. McGraw-Hill Professional, 1994.
- [53] Cedric Langbort. A mechanism design approach to dynamic price-based control of multi-agent systems. In Rolf Johansson and Anders Rantzer, editors, *Distributed Decision Making and Control*, volume 417 of *Lecture Notes in Control and Information Sciences*, pages 113–129. Springer Berlin / Heidelberg, 2012.
- [54] Jason Lepore. Cournot outcomes under Bertrand-Edgeworth competition with demand uncertainty.

*Journal of Mathematical Economics*, 48(3):177–186, 2012.

- [55] F.L. Lewis and V.L. Syrmos. *Optimal Control*. A Wiley-Interscience publication. John Wiley & Sons, 1995.
- [56] A. Mas-Colell, M.D. Whinston, and J.R. Green. *Microeconomic theory*. Oxford University Press, 1995.
- [57] Barry A Mather. Quasi-static time-series test feeder for PV integration analysis on distribution systems. In *Power and Energy Society General Meeting, 2012 IEEE*, pages 1–8. IEEE, 2012.
- [58] Johanna L. Mathieu, Ashok J. Gadgil, Duncan S. Callaway, Phillip N. Price, and Sila Kiliccote. Characterizing the response of commercial and industrial facilities to dynamic pricing signals from the utility. In *In Proceedings of ASME 4th International Conference on Energy Sustainability*, 2010.
- [59] Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V. Vazirani. *Algorithmic Game Theory*. Cambridge University Press, New York, NY, USA, 2007.
- [60] M.J. Osborne and A. Rubinstein. *A course in game theory*. MIT Press, 1994.
- [61] Jukka V Paatero and Peter D Lund. Effects of large-scale photovoltaic power integration on electricity distribution networks. *Renewable Energy*, 32(2):216–234, 2007.
- [62] S. B. Peterson, J. F. Whitacre, and J. Apt. The economics of using plug-in hybrid electric vehicle battery packs for grid storage. *Journal of Power Sources*, 195:2377–2384, 2010.
- [63] A.G. Phadke. Synchronized phasor measurements in power systems. *Computer Applications in Power, IEEE*, 6(2):10–15, Apr. 1993.
- [64] Stanley S. Reynolds and Bart J. Wilson. Bertrand-Edgeworth competition, demand uncertainty, and asymmetric outcomes. *Journal of Economic Theory*, 92(1):122 – 141, 2000.
- [65] Michael H. Rothkopf. Thirteen reasons why the Vickrey-Clarke-Groves process is not practical. *Operations Research*, 55(2):191–197, Apr. 2007.
- [66] P. W. Sauer and M. A. Pai. *Power system dynamics and stability*. Stipes Publishing Co., 1998.
- [67] Kevin P Schneider, Yousu Chen, David P Chassin, Robert G Pratt, David W Engel, and Sandra Thompson. Modern grid initiative: Distribution taxonomy final report. Technical report, Pacific Northwest National Laboratory, 2008.
- [68] Kevin P Schneider, J Fuller, F Tuffner, and Ruchi Singh. Evaluation of conservation voltage reduction (CVR) on a national level. Technical report, Pacific Northwest National Laboratory report, 2010.
- [69] Fred C. Schweppe, Michael C. Caramanis, Richard D. Tabors, and Roger E. Bohn. *Spot pricing of electricity*. Kluwer Academic Publishers, Boston, MA, 1988.
- [70] M. Shahidehpour, H. Yamin, and Z. Li. *Market Operations in Electric Power Systems (Forecasting, Scheduling, and Risk Management)*. Wiley-IEEE Press, 2002.
- [71] Raymond R Shoults and Larry D Swift. Power system loads. In Leonard L Grigsby, editor, *Electric Power Generation, Transmission, and Distribution*, chapter 20. CRC Press, 2012.

- [72] Daniel S Shugar. Photovoltaics in the utility distribution system: the evaluation of system and distributed benefits. In *Photovoltaic Specialists Conference, 1990., Conference Record of the Twenty First IEEE*, pages 836–843. IEEE, 1990.
- [73] C. L. Su and D. Kirschen. Quantifying the effect of demand response on electricity markets. *IEEE Transactions on Power Systems*, 24(3):1199–1207, 2009.
- [74] Mariano Ventosa, Álvaro Bañalao, Andrés Ramos, and Michel Rivier. Electricity market modeling trends. *Energy Policy*, 33(7):897 – 913, 2005.
- [75] V. Vittal, J. McCalley, V. Ajjarapu, and U.V. Shanbhag. Impact of increased DFIG wind penetration on power systems and markets. Technical Report PSERC 09-10, Power Systems Engineering Research Center, 2009.
- [76] A. J. Wood and B. F. Wollenberg. *Power generation, operation, and control*. Knovel, second edition, 1996.
- [77] Achim Woyte, Vu Van Thong, Ronnie Belmans, and Johan Nijs. Voltage fluctuations on distribution level introduced by photovoltaic systems. *Energy Conversion, IEEE Transactions on*, 21(1):202–209, 2006.
- [78] Tong Wu, M. Rothleder, Z. Alaywan, and A.D. Papalexopoulos. Pricing energy and ancillary services in integrated market systems by an optimal power flow. *Power Systems, IEEE Transactions on*, 19(1):339 – 347, Feb. 2004.
- [79] Z. Yu, F.T. Sparrow, T.L. Morin, and G. Nderitu. A Stackelberg price leadership model with application to deregulated electricity markets. In *Power Engineering Society Winter Meeting, 2000. IEEE*, volume 3, pages 1814 –1819, Jan. 2000.
- [80] Tongxin Zheng and E. Litvinov. Ex post pricing in the co-optimized energy and reserve market. *Power Systems, IEEE Transactions on*, 21(4):1528 –1538, Nov. 2006.
- [81] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas. MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education. *Power Systems, IEEE Transactions on*, PP(99):1 –8, June 2010.

## 4.4 Models and control strategies for optimal control of aggregated FirmPV systems

### Summary

This is one of three final reports completed by a UC Berkeley team of students and faculty in the Energy and Resources Group. It covers efforts from Task 4.4. Our specific focus is on understanding distribution system impacts and the opportunity for creating value by incorporating storage into the CAISO's dispatch process.

This particular report describes our efforts to understand how to value and control storage in large-scale power systems. The report is written in two parts.

In Part 1 (Section 4.4.1), we develop control strategies for energy storage to absorb renewable generator production variability. The control strategies are based on storage either cooperating directly with a renewable producer, or multiple storage units competing in a more general market environment. We frame the optimal storage scheduling strategy as an inventory control-like policy. By making a variety of mild approximations (e.g. neglecting storage round trip efficiency, or certain constraints on storage operation) we show that the inventory control policy can be represented as very simple linear relationships, easily computed in real time. Due to the approximations, the control policies are suboptimal, but numerical simulations indicate that the suboptimality is on the order of a few percent.

Part 2 (Section 4.4.2) studies the allocation of storage resources – in time, space, and across market products – to evaluate the economic value of energy storage at the wholesale level. While many possible functions for grid-connected storage have been identified, the overall value of storage performing one or more of these functions in a networked power system has not been fully characterized. We present a DC load flow model that enables storage devices to perform any of the following grid functions: energy arbitrage, regulation, and load following. The model is based on a 240-bus model that is representative of the Western Interconnection grid, and includes coal, gas, nuclear, biomass, geothermal, hydro, solar, and wind generation. As storage penetration is increased we find that the system operating cost is reduced. This reduction is primarily due to storage being able to provide regulation and load following, and a reduction in the number of generator starts. We also find that under some system conditions increased storage penetration could increase the carbon intensity of the system.



#### 4.4.1 Competitive energy storage in the presence of renewables

**4.4.1.1 Introduction** Distributed generators [1], roughly defined as producers of small quantities of power geographically close to the loads they serve, are becoming increasingly common. The advantages of distributed generation are numerous, including microgrid operation [39], enhanced robustness via dispersion of vulnerabilities, reduced substation capacity requirements, and efficient transmission of power due to shorter power flow paths [7, 61]. One of the key obstacles to the integration of distributed wind and solar generators is intermittency: the outputs of wind and solar producers are variable and unpredictable, and require additional flexible resources (or operating reserves) to maintain the supply-demand balance [45].

By co-locating energy storage with distributed generators, intermittency can be balanced locally [8, 20, 35], in turn reducing requirements for transmission infrastructure and operating reserves from the bulk grid. When storage and generation are under the same control, enhanced performance is achievable [24]. However, in restructured or deregulated markets, independent storage agents could enter the market as well. In this case, similar benefits are attainable, but, because generation and storage are not operated by a single entity, the parties involved must agree on some economic terms.

Existing reserve and ancillary service markets capture such economic terms for generators [4, 63], in which a premium is paid for reserve capacity, which may then be purchased at some strike price; we will henceforth use the phrase reserve market. Storage, on the other hand, exists primarily as a reserve, and so does not incur significant opportunity costs [30], and in that standard storage technologies such as pumped hydro [16] and batteries [51] must regularly accept external power to meet their purposes. Toward enabling energy storage to provide reserves alongside generators, independent system operators are considering mechanisms for storage to participate in reserve markets [18]. In such setups payments to storage will be a combination of capacity, mileage, and energy payments [17]; here we focus on the latter as a profitable channel. Note that in the context of renewable production, such payments may also be interpreted as imbalance penalties, and are implemented in various existing markets [9, 15, 48, 59].

In this work, we examine how independent operators of storage might derive profits from absorbing unanticipated renewable variability in reserve markets, and the resulting interaction between storage and a renewable energy producer. Implicit in our framework is a secondary profitable channel: inter-temporal price arbitrage. We remark that with the exception of pumped hydro, the latter has not been pursued in practice, because storage inefficiencies and capital costs have outweighed expected profits even under high price volatility [14]. However, the two objectives are seamlessly integrated in our framework, so that any amount of profitable arbitrage will be identified and balanced with absorbing unexpected variability.

In our main result, we study how storage can pursue these objectives through scheduled transactions when it is paid through capacity premiums and imbalance fees in a reserve market. When energy imbalances are temporally uncorrelated and energy transactions are perfectly efficient, we find that the optimal storage cost and scheduling policy are piecewise-affine in the current energy level, and are thus easy to compute and analyze. We build on this result in the correlated and inefficient cases to construct heuristic solutions with similar tractability, but which well-approximate the general optimal solution. By assuming temporal uniformity of parameters and perfect injection and extraction efficiencies, we derive a simplified single period approximation, which we use to study the Nash equilibrium between storage and renewable energy producers. We find that storage inefficiency can lead to non-socially optimal equilibria.

Our work builds on inventory control theory [5, 6, 10, 11, 49, 55], which has been used to optimize generator reserve management [21] and as well as energy storage formulations related to our own. In [37], energy pre-commitments are optimized assuming uniformly distributed energy imbalances, and closed-form, infinite horizon policies are used to estimate optimal sizing. In [41], approximate dynamic programming algorithms are used to solve a similar formulation. In [64], heuristic policies are developed for a version of the problem without injection and extraction inefficiencies. In each of these formulations, it is assumed that energy storage is operated alongside a renewable producer primarily to hedge against uncertainty. We comment that the related problem of optimally contracting wind, which has been shown to be a single-period newsvendor problem [15, 48], is the non-dynamic analog of our problem.

Our formulation applies to these cooperative scenarios, but also to independent energy storage, which we further explore in a game-theoretic analysis. As with existing literature, the scheduling aspect of our work applies most naturally in five minute markets, in which the policy can be applied to the current state, but can also be used to schedule transactions in hour- or day-ahead markets [37]. By definition, the variability absorption component takes place in reserve markets. Our work extends existing results from traditional inventory control by modeling details particular to energy storage, such as physical inefficiencies and hybrid storage configurations [60].

The section of the report is outlined as follows. In Section 4.4.1.2, we define our notation and introduce a basic framework for pricing deviations from the producer covered by storage. We then optimize storage operation in Section 4.4.1.3, and discuss the two market contexts our model addresses. We then look at the equilibrium between a renewable producer and storage in Section 4.4.1.4, and conclude in Section 4.4.1.5 with numerical examples.

**4.4.1.2 Modeling** In this section, we describe a generic model of how storage may profit either from traditional capacity premiums or fees assessed on a renewable producer’s contract deviations. These fees are sometimes referred to as imbalance prices, and differ from market to market [48]; for example, in Spain, they are proportional to the market price [59]. In this work, we will assume that they are known functions of the energy imbalances. The resulting financial exchange is similar to that considered for a single wind producer, in which some market price is paid for producer’s contracted production, and a fee is assessed on deviations [9, 15, 48]. In this work we adopt the perspective of the agent assessing the fee.

**4.4.1.2.1 Setup and notation** Two types of entities are distinguished by their economic behavior: storage, which extracts profits by absorbing deviations, and all other agents, such as generators and loads. For example, a load provides a forecast, similar to that of a variable producer but negative, and is penalized for deviations. Variable production of energy,  $w$  is divided into a scheduled (ex-ante) component and a deviation,  $w = \bar{w} + \hat{w}$ . The energy flow into and out of storage, which is fully controllable and maintains the energy balance, is also divisible into scheduled and balancing components,  $\bar{s}$  and  $\hat{s}$ , and current holdings are denoted by  $s$ . Note that energy flows into and out of storage refer to grid-side quantities, and scalar efficiencies are used to convert to internally stored energy. Table 1 summarizes the variable quantities.

Scheduled transactions need not balance, e.g.  $\bar{w} - \bar{s}$  may be equal to a deterministic load or transaction with the transmission system. We do however assume that deviations balance, so that  $\hat{w} = \hat{s}$ ; this is made feasible in our framework by defining the transmission system as a storage option with unlimited capacity and holding. While in many ways it is incorrect to view the bulk grid as a storage entity, it serves

Table 1: Variables

Symbol	Description
$s$	Stored potential energy
$\hat{s}$	Unscheduled energy transaction requested of storage
$\bar{s}$	Scheduled energy transaction of storage
$\hat{w}$	Unscheduled energy production from producer
$\bar{w}$	Scheduled energy production from producer

our present purpose by ensuring that all energy is balanced, albeit at a possibly large cost.

$(\cdot)^+ = \max\{\cdot, 0\}$  is used to denote the argument if it is positive and zero otherwise, and likewise for  $(\cdot)^- = \min\{\cdot, 0\}$ ; parentheses are omitted when the argument is a single term. We use the convention that a storage “injection” is delivery of electricity to the grid, and “extraction” is removal from the grid. We note that storage cannot simultaneously inject and extract power. Throughout, injection and extraction quantities are respectively denoted by superscripts  $i$  and  $e$ .

Table 2: Defining parameters

Symbol	Description
$c_t$	Ex-ante price for all transactions at time $t$
$\delta_{k,t}^i, \delta_{k,t}^e$	Ex-post imbalance fee functions of storage $k$ at $t$
$\eta_{k,t}^i, \eta_{k,t}^e$	Capacity premium functions of storage $k$ at $t$
$\lambda_k^i, \lambda_k^e \in [0, 1]$	Efficiencies of storage $k$
$\kappa_k^i \leq 0, \kappa_k^e \geq 0$	Discharge and charge rate limit of storage $k$
$\alpha_k \in [0, 1]$	Leakage coefficient of storage $k$
$S_k$	Energy capacity of storage $k$
$\theta$	Energy imbalance intertemporal correlation

The parameters in Table 2 specify our framework.  $\bar{w}$  is chosen by the renewable producer, and  $\bar{s}$  by the storage facility.  $c$ ,  $\delta^i$ ,  $\delta^e$ ,  $\eta^i$ , and  $\eta^e$  are exogenous to our model; although we assume deterministic prices here, we comment on how one might incorporate uncertainty in Section 4.4.1.3.3. We assume that  $c_t \geq 0$ , which renders simultaneous injection and extraction uniformly unprofitable; we remark that our subsequent developments are still applicable if this assumption is dropped, but optimality may be lost. The capacity premiums  $\eta^i$  and  $\eta^e$  are generic functions of the injection and extraction capacities. The imbalance fees  $\delta^i$  and  $\delta^e$  are generic functions of  $\hat{s}_t$  that may depend on a variety of factors, for example the magnitude of  $\hat{s}_t$ , spot prices [54] and congestion rents [31] arising from network location, and fixed costs [38]. Note that while they are more naturally interpreted as positive quantities, our formulation allows them to be positive or negative. We make the approximation that  $\kappa^i$  and  $\kappa^e$  are constants, i.e. are independent of the current state; in Section 4.4.1.3.2 we discuss when this assumption may be removed. Note that transmission MW capacity constraints may also be encoded in  $\kappa^i$  and  $\kappa^e$ . We comment that the leakage parameter,  $\alpha$  might be very close to one for all but long time-scale applications; however, since it is not a significant source of model complexity, we include it for completeness.

**4.4.1.2.2 Utility functions** Suppose a collection of renewable sources, subscripted by  $w$ , produce a certain amount of variability that is absorbed by  $n_s$  storages, and that each storage  $k$  receives  $\hat{s}_k$  of the total variability. Let  $S_k^i$  and  $S_k^e$  be the up and down capacity committed by storage  $k$ ; we will discuss the determination of these quantities later. The following utility functions express the financial involvement of

each player at a given time; because time indices are identical, we suppress them here.

$$u_w^j = c(\bar{w}_j + \hat{w}_j) + \frac{1}{n_s} \left[ \sum_k \delta_k^i(\hat{s}_k^-) - \delta_k^e(\hat{s}_k^+) \right] \quad (1)$$

$$u_s^k = -c(\bar{s}_k + \hat{s}_k) - \delta_k^i(\hat{s}_k^-) + \delta_k^e(\hat{s}_k^+) + \eta_k^i(S_k^i) + \eta_k^e(S_k^e) \quad (2)$$

**4.4.1.2.3 Allocating deviations to storage** We now describe how energy is divided amongst each storage unit given their current holding and total capacity. Deviations are absorbed in the least expensive way possible for a prescribed set of fees. This can be expressed via the following optimization:

$$\begin{aligned} \min_{s_k^e, s_k^i} \quad & \sum_k \delta_k^e(s_k^e) - \delta_k^i(s_k^i) \\ \text{s. t.} \quad & \sum_k s_k^e - s_k^i = \sum_j \hat{w}_j \\ & s_k^e \geq 0, s_k^i \geq 0 \\ & 0 \leq \alpha_k s_k + (\bar{s}_k^- - s_k^i) / \lambda_k^i + \lambda_k^e (\bar{s}_k^+ + s_k^e) \leq S_k \\ & \lambda_k^i \kappa_k^i \leq \bar{s}_k + s_k^e - s_k^i \leq \kappa_k^e / \lambda_k^e \end{aligned}$$

where  $s_k$  is the energy holding of storage  $k$  at the previous time stage. The ex-post balancing transaction is then  $\hat{s}_k = s_k^e - s_k^i$ . For the remainder of this work, we will only consider a single source of variability, and so will not concern ourselves with dividing penalties fairly among multiple producers. In Section 4.4.1.4, we model the transmission system as the most expensive pair  $\delta^i$  and  $\delta^e$ , with unlimited capacity,  $S$ , and holdings,  $s$ .

Since ex-ante transactions  $\bar{w}$  and  $\bar{s}$  need not balance each other, the producer's primary interest is selling as much power as possible at market price while minimizing deviation fees, and that of storage is maintaining its holdings so as to best absorb deviations, and arbitraging over temporal market price variations.

**4.4.1.2.4 Market context** We now briefly discuss how the above formulations apply in various market scenarios.

**Cooperation with a variable producer** Suppose a single renewable producer is storing energy to augment their own production profits. In renewable operation, one may determine the contracted production, set in either day- or hour-ahead markets [15, 48], but there are fewer means to control power output compared to conventional generation. Such contracts may be optimized taking into account available energy storage [13], leaving the remaining operational storage decisions to future times or finer time scales.

Assume that  $\delta_t^i(\hat{s}_t)$  and  $\delta_t^e(\hat{s}_t)$  are the total energy imbalance costs faced by the producer. If the storage is owned by the producer, storage utility in (2) exactly aligns with producer interests by capturing imbalance costs while also arbitraging over the market price, i.e. maximizing producer profits. Because the storage and producer cooperate in this case, the producer's estimate of  $\hat{w}$  can be made directly available to the storage (this is in contrast to case (2) below, in which storage may not have a direct estimate of  $\hat{w}$ ). Since it is not bidding capacity directly into a market,  $\eta^i(x) = \eta^e(x) = 0$ .

**Independent operation** Now suppose that an independent storage is attempting to maximize its profits. We assume that storage arbitrages energy via injections and extractions in forward energy markets (day- or hour-ahead), and dedicates its remaining capacity to a reserve or ancillary service market. It is paid in the latter on the basis of energy through  $\delta^i$  and  $\delta^e$  and/or capacity through  $\eta^i$  and  $\eta^e$ .

While the current practice for procuring generation resources for ancillary services is capacity-based, arrangements that include additional payments for energy imbalances are currently under consideration by independent system operators. An example is Regulation Energy Management of the California Independent System Operator, in which the quantities  $\delta_t^i$  and  $\delta_t^e$  model energy mileage payments [17, 18]. Reasons for paying storage on the basis of energy and not exclusively capacity include that capacity payments do not capture losses due to inefficiencies and the contribution of repeated cycling to lifetime costs [35]. In contrast to the previous market scenario, in this case the energy imbalance probability distribution may be provided by a market operator or empirically estimated by the storage operator.

We remark that the assumption that storage is a cheaper option than generator reserves is specific to the current context, and that some storage may be more expensive. In this respect, the work is in anticipation of scenarios in which storage does represent less expensive options, or is preferred for superior performance characteristics [25, 47].

**4.4.1.3 Optimal ex-ante scheduling** We now formulate the dynamic optimization problem arising from selecting  $\bar{s}_t$  at each time stage  $t = 1, \dots, T - 1$  to maximize profits from absorbing deviations. As stated in the introduction, we will find that the optimal scheduling problem falls within the domain of inventory control. In our development, we follow the derivations of Ch. 4, Section 2 of [12] and Section 2 of [46], the latter of which provides a thorough survey of inventory control results.

Consider a single storage unit, with distributions for unscheduled injections and extractions given by the first order autoregressive process

$$\hat{s}_{t+1} = \theta \hat{s}_t + \epsilon_{t+1}, \quad (3)$$

where  $\theta \in [0, 1]$  and  $\epsilon_t$  is drawn from the distribution  $f_t(\epsilon_t)$  (with cumulative distribution function  $F_t$ ). We comment that the distribution  $f_t(\epsilon_t)$  is related to the net system imbalance and its allocation to storage; we consider these dependencies in the next section, but for now assume that the net uncertainty presented to a single storage is described by a single, known distribution.

We assume that beyond its forward market decisions (at prices  $c_t$ ), storage chooses to commit the remainder of its capacity to an ancillary service market, for which it is paid according to the capacity premiums  $\eta^i$  and  $\eta^e$  and imbalance fees  $\delta^i$  and  $\delta^e$ .

One relaxation is made between here and the previous section: the charge rate constraint is applied separately to ex-ante transactions,  $\bar{s}$ , and deviations,  $\hat{s}$ . While it is pragmatic to constrain the scheduled power in this way prior to the deviation, it is the sum of scheduled power and the deviation that should be constrained as such in the actual physical exchange. However, on larger time scales, e.g. one hour, it may be more realistic to model ex-ante transactions and deviations as occurring asynchronously, in which case jointly constraining them will introduce an extra degree of conservatism.

**4.4.1.3.1 Stochastic dynamic optimization** We encode capacity and charge rate limits in a saturation function, defined as

$$\text{sat}(x_1, x_2) = \max \left\{ \min \{x_1, \min \{S - x_2, \kappa^e\}\}, \max \{-x_2, \kappa^i\} \right\}.$$

This truncates a change in stored energy  $x_1$  at storage capacity and charge and discharge rate limits, which are determined by the stored energy,  $x_2$ . The storage holdings state,  $s_t$ , has dynamics (see, e.g. [38]) given by

$$s_{t+1} = \alpha s_t + \bar{s}_t^- / \lambda^i + \lambda^e \bar{s}_t^+ + \text{sat}(\hat{s}_t^- / \lambda^i + \lambda^e \hat{s}_t^+, \alpha s_t + \bar{s}_t^- / \lambda^i + \lambda^e \bar{s}_t^+)$$

The purpose of the saturation function is to constrain the absorption of imbalances to within the physical capabilities of the storage device. Define the function

$$g_t(x_1, x_2) = -c_t \lambda^i \text{sat}(x_1^- / \lambda^i, x_2) - \delta_t^i (\lambda^i \text{sat}(x_1^- / \lambda^i, x_2)) - c_t \text{sat}(\lambda^e x_1^+, x_2) / \lambda^e + \delta_t^e (\text{sat}(\lambda^e x_1^+, x_2) / \lambda^e).$$

Then expected profits from deviations are given by

$$\bar{g}_t(\hat{s}_{t-1}, x) = \mathbb{E} [g_t(\theta \hat{s}_{t-1} + \epsilon_t, x)].$$

In all but special cases, this and further expectations in this section must be evaluated numerically, e.g. using quadrature or Monte Carlo integration [34]; since all integrals are one dimensional, it does not add substantial computational burden.

The committed up and down capacities are given by  $\min \{\alpha s_t + \bar{s}_t^- / \lambda^i + \lambda^e \bar{s}_t^+, -\kappa^i\}$  and  $\min \{S - \alpha s_t - \bar{s}_t^- / \lambda^i - \lambda^e \bar{s}_t^+, \kappa^e\}$ , respectively.

We seek to solve the following optimization:

$$\begin{aligned} \max_{\bar{s}_t} \quad & g_T(s_T) + \sum_{t=0}^{T-1} \bar{g}_t(\hat{s}_{t-1}, \alpha s_t + \bar{s}_t^- / \lambda^i + \lambda^e \bar{s}_t^+) - c_t \bar{s}_t \\ & + \eta_t^i (\lambda^i \min \{\alpha s_t + \bar{s}_t^- / \lambda^i + \lambda^e \bar{s}_t^+, -\kappa^i\}) \\ & + \eta_t^e (\min \{S - \alpha s_t - \bar{s}_t^- / \lambda^i - \lambda^e \bar{s}_t^+, \kappa^e\} / \lambda^e) \\ \text{s. t.} \quad & 0 \leq \alpha s_t + \bar{s}_t^- / \lambda^i + \lambda^e \bar{s}_t^+ \leq S \\ & \lambda^i \kappa^i \leq \bar{s}_t \leq \kappa^e / \lambda^e \end{aligned}$$

Note that the above does not contain the “dynamic” constraint; the desired optimization is conveyed exactly in the dynamic programming formulation below, which we use in the remainder of our

analysis.

$$\begin{aligned}
J_T^1(s_T) &= g_T(s_T) \\
J_t^1(\hat{s}_{t-1}, s_t) &= \max_{\bar{s}_t} \bar{g}_t(\hat{s}_{t-1}, \alpha s_t + \bar{s}_t^- / \lambda^i + \lambda^e \bar{s}_t^+) - c_t \bar{s}_t \\
&\quad + \eta_t^i (\lambda^i \min \{ \alpha s_t + \bar{s}_t^- / \lambda^i + \lambda^e \bar{s}_t^+, -\kappa^i \}) \\
&\quad + \eta_t^e (\min \{ S - \alpha s_t - \bar{s}_t^- / \lambda^i - \lambda^e \bar{s}_t^+, \kappa^e \} / \lambda^e) \\
&\quad + \mathbb{E} \left[ J_{t+1}^1 \left( \theta \hat{s}_{t-1} + \epsilon_t, \alpha s_t + \bar{s}_t^- / \lambda^i + \lambda^e \bar{s}_t^+ \right. \right. \\
&\quad \left. \left. + \text{sat} \left( (\theta \hat{s}_{t-1} + \epsilon_t)^- / \lambda^i + \lambda^e (\theta \hat{s}_{t-1} + \epsilon_t)^+ , \right. \right. \right. \\
&\quad \left. \left. \left. \alpha s_t + \bar{s}_t^- / \lambda^i + \lambda^e \bar{s}_t^+ \right) \right) \right] \\
\text{s. t.} \quad &0 \leq \alpha s_t + \bar{s}_t^- / \lambda^i + \lambda^e \bar{s}_t^+ \leq S \\
&\lambda^i \kappa^i \leq \bar{s}_t \leq \kappa^e / \lambda^e
\end{aligned}$$

**4.4.1.3.2 Inventory control substitution** The main technical mechanism taken from inventory control is the pair of substitutions  $y^i = \alpha s_t + \bar{s}_t^- / \lambda^i$  and  $y^e = \alpha s_t + \lambda^e \bar{s}_t^+$ . Throughout, we will use the letter  $y$  to indicate the post-control, pre-imbalance stored energy when it is a quantity being optimized, and  $z_t$  after it has been optimized; because  $y$  is an optimization variable, it is not subscripted. Define the constraint sets

$$\begin{aligned}
\beta^i(s_t) &= \{x \mid \max \{0, \kappa^i + \alpha s_t\} \leq x \leq \alpha s_t\} \\
\beta^e(s_t) &= \{x \mid \alpha s_t \leq x \leq \min \{S, \kappa^e + \alpha s_t\}\}
\end{aligned}$$

and the functions

$$\begin{aligned}
G_t^i(\hat{s}_{t-1}, y^i) &= \bar{g}_t(\hat{s}_{t-1}, y^i) - c_t \lambda^i y^i \\
&\quad + \eta_t^i (\lambda^i \min \{y^i, -\kappa^i\}) \\
&\quad + \mathbb{E} \left[ J_{t+1}^1(y^i + \text{sat}((\theta \hat{s}_{t-1} + \epsilon_t)^- / \lambda^i \right. \\
&\quad \left. + \lambda^e (\theta \hat{s}_{t-1} + \epsilon_t)^+, y^i)) \right] \tag{4}
\end{aligned}$$

$$\begin{aligned}
G_t^e(\hat{s}_{t-1}, y^e) &= \bar{g}_t(\hat{s}_{t-1}, y^e) - c_t y^e / \lambda^e \\
&\quad + \eta_t^e (\min \{S - y^e, \kappa^e\} / \lambda^e) \\
&\quad + \mathbb{E} \left[ J_{t+1}^1(y^e + \text{sat}((\theta \hat{s}_{t-1} + \epsilon_t)^- / \lambda^i \right. \\
&\quad \left. + \lambda^e (\theta \hat{s}_{t-1} + \epsilon_t)^+, y^e)) \right]. \tag{5}
\end{aligned}$$

It is now evident why these substitutions would not be valid if charge rate constraints were imposed on the state evolution: an isolated  $\alpha s_t$  term would appear in the argument of  $J_{t+1}$ , and  $G_t^i$  and  $G_t^e$  would not solely be functions of  $y^i$  and  $y^e$ . The dynamic programming equation is equivalent to

$$\begin{aligned}
J_t^1(\hat{s}_{t-1}, s_t) &= \max \left\{ c_t \alpha \lambda^i s_t + \max_{y^i \in \beta^i(s_t)} G_t^i(\hat{s}_{t-1}, y^i), \right. \\
&\quad \left. c_t \alpha s_t / \lambda^e + \max_{y^e \in \beta^e(s_t)} G_t^e(\hat{s}_{t-1}, y^e) \right\}. \tag{6}
\end{aligned}$$

Notice that in (6), storage may not simultaneously extract and inject power, which it was able to do in the previous dynamic programming formulation. The two formulations are equivalent because we have

assumed that  $c_t \geq 0$ ; without this assumption, (6) and subsequent policies would be applicable, but not optimal, because it would not capture situations in which it is profitable to dissipate energy. We comment however that energy dissipation may only be practical in certain types of storage, for example by releasing reservoir water through a spillway as opposed to running current through a battery, the latter of which would degrade state of health.

Because the maximum of two concave functions in general is not concave, it is difficult to say much about  $G_t^i$  and  $G_t^e$ , even if  $\bar{g}_t$  does have desirable properties. However, because each maximization is a bounded function on a compact or finite set, all maxima are attained. Define

$$\begin{aligned} z_t^i(\hat{s}_{t-1}, s_t) &= \underset{y^i}{\operatorname{argmax}} \{ G_t^i(\hat{s}_{t-1}, y^i) \mid y^i \in \beta^i(s_t) \} \\ z_t^e(\hat{s}_{t-1}, s_t) &= \underset{y^e}{\operatorname{argmax}} \{ G_t^e(\hat{s}_{t-1}, y^e) \mid y^e \in \beta^e(s_t) \}. \end{aligned} \quad (7)$$

The general-case optimal policy and cost-to-go is given below.

**Policy 1 (Optimal)** *The optimal energy transaction at time  $t$  is given by*

$$\mu_t^1(\hat{s}_{t-1}, s_t) = \begin{cases} \lambda^i(z_t^i(\hat{s}_{t-1}, s_t) - \alpha s_t) \text{ if } c_t \alpha (\lambda^i - 1/\lambda^e) s_t \\ \geq G_t^e(\hat{s}_{t-1}, z_t^e(\hat{s}_{t-1}, s_t)) - G_t^i(\hat{s}_{t-1}, z_t^i(\hat{s}_{t-1}, s_t)) \\ (z_t^e(\hat{s}_{t-1}, s_t) - \alpha s_t)/\lambda^e \text{ if } c_t \alpha (\lambda^i - 1/\lambda^e) s_t \\ \leq G_t^e(\hat{s}_{t-1}, z_t^e(\hat{s}_{t-1}, s_t)) - G_t^i(\hat{s}_{t-1}, z_t^i(\hat{s}_{t-1}, s_t)), \end{cases}$$

*and the associated cost by*

$$J_t^1(\hat{s}_{t-1}, s_t) = \begin{cases} c_t \lambda^i \alpha s_t + G_t^i(\hat{s}_{t-1}, z_t^i(\hat{s}_{t-1}, s_t)) \text{ if } c_t \alpha (\lambda^i - 1/\lambda^e) s_t \\ \geq G_t^e(\hat{s}_{t-1}, z_t^e(\hat{s}_{t-1}, s_t)) - G_t^i(\hat{s}_{t-1}, z_t^i(\hat{s}_{t-1}, s_t)) \\ c_t \alpha s_t / \lambda^e + G_t^e(\hat{s}_{t-1}, z_t^e(\hat{s}_{t-1}, s_t)) \text{ if } c_t \alpha (\lambda^i - 1/\lambda^e) s_t \\ \leq G_t^e(\hat{s}_{t-1}, z_t^e(\hat{s}_{t-1}, s_t)) - G_t^i(\hat{s}_{t-1}, z_t^i(\hat{s}_{t-1}, s_t)). \end{cases}$$

$z_t^i$  and  $z_t^e$  are optimal set-points, which the control respectively tracks depending on whether an injection or extraction is performed.  $\mu^1$  and  $J^1$  may be computed by discretizing both states, per usual in dynamic programming. However, by proceeding in this way we make no use of the inventory control structure inherent in this problem. We remark that although the discussed purpose is to profit from deviations in productions, temporal arbitrage is built in to the optimal policy; indeed,  $c_t$  through  $c_T$  all appear in  $G_t^i$  and  $G_t^e$ , and may exert influence if prices are volatile. Finally, in the absence of capacity payments, the inventory control substitutions and the developments of the next subsection would remain valid even if charge and discharge rate limits were functions of the state,  $\kappa^e(s_t)$  and  $\kappa^i(s_t)$ .

**4.4.1.3.3 Suboptimal policies** We now focus our attention on suboptimal policies, which are easier to compute and straightforward to derive via approximating **Policy 1**. Specifically, we seek policies that are affine in either  $s_t$  or  $\hat{s}_t$ .



**Approximate efficiency and rate limits** **Policy 1** has complex dependence on  $s_t$  due to charge rate constraints and inefficiencies, the latter of which induce hybrid dynamics. Notice that the only difference between  $G_t^i$  and  $G_t^e$  is their second term; replacing this term with  $-c_t y$  allows us to circumvent the hybrid aspect of the dynamics. In this case, both are equivalent to

$$\begin{aligned} G_t(\hat{s}_{t-1}, y) &= \bar{g}_t(\hat{s}_{t-1}, y) - c_t y \\ &\quad + \mathbb{E} \left[ J_{t+1}(y + \text{sat}((\theta \hat{s}_{t-1} + \epsilon_t)^- / \lambda^i \right. \\ &\quad \left. + \lambda^e (\theta \hat{s}_{t-1} + \epsilon_t)^+, y)) \right]. \end{aligned} \quad (8)$$

Let

$$z_t(\hat{s}_{t-1}) = \underset{0 \leq y \leq S}{\operatorname{argmax}} G_t(\hat{s}_{t-1}, y). \quad (9)$$

Similarly modifying  $J_t^1$ , the approximate cost to go is then

$$J_t^2(\hat{s}_{t-1}, s_t) = c_t \alpha s_t + G_t(\hat{s}_{t-1}, z_t(\hat{s}_{t-1}))$$

$J_t^2$  is affine in  $s_t$ , the most significant implication of which is that  $s_t$  does not need to be discretized. We must still account for charge rate constraints, however, preferably without sacrificing the piecewise affine structure. We apply them later by defining

$$\tilde{z}_t(z_t(\hat{s}_{t-1}), s_t) = \max \left\{ \min \{ z_t(\hat{s}_{t-1}), \alpha s_t + \kappa^e \}, \alpha s_t + \kappa^i \right\}.$$

We then have the following suboptimal policy:

**Policy 2 (Approximate efficiency and rate limits)**

$$\begin{aligned} \mu_t^2(\hat{s}_{t-1}, s_t) &= \lambda^i (\tilde{z}_t(z_t(\hat{s}_{t-1}), s_t) - \alpha s_t)^- \\ &\quad + (\tilde{z}_t(z_t(\hat{s}_{t-1}), s_t) - \alpha s_t)^+ / \lambda^e \end{aligned}$$

Only  $\hat{s}_t$  needs discretization to find  $\mu_t^2$ , and  $s_t$  may be treated continuously. In effect, we have made a mild approximation to efficiencies, and are treating the charge rate limits myopically. We remark that just setting  $\lambda^i$  and  $\lambda^e$  to one may result in infeasible energy transactions.

**Small correlations** When  $\theta$  is small, we may approximate the energy imbalances as temporally uncorrelated, so that  $\hat{s}_t = \epsilon_t$ , and we obtain the following suboptimal policy.

**Policy 3 (Small correlations)** In **Policy 1**, replace all instances of  $\hat{s}_{t-1}$  with zero.

Note that in this policy, one could apply the assumptions of **Policy 2** to obtain a purely affine policy, essentially recovering the standard inventory control result of [12, 46].

**Small arbitrage** We now consider the case that profits from arbitrage are small relative to those from imbalances, for example such that

$$\max_{t_1, t_2 \in \{1, \dots, T\}} |c_{t_1} - c_{t_2}| \ll \min_t \min \{ \delta_t^i, \delta_t^e \}.$$

Because profits will predominantly come from absorbing variability, we incorporate the nonrandom components of  $\hat{s}_t, \theta\hat{s}_{t-1}$ , into the constraints on  $\bar{s}_t$ , as shown in the dynamic program below:

$$\begin{aligned}
J_T^4(\hat{s}_{T-1}, s_T) &= g_T(\hat{s}_{T-1}, s_T) \\
J_t^4(\hat{s}_{t-1}, s_t) &= \max_{\bar{s}_t} \bar{g}_t(0, \alpha s_t + (\bar{s}_t + \theta\hat{s}_{t-1})^- / \lambda^i \\
&\quad + \lambda^e(\bar{s}_t + \theta\hat{s}_{t-1})^+ - c_t(\bar{s}_t + \theta\hat{s}_{t-1}) \\
&\quad + \eta_t^i (\lambda^i \min \{ \alpha s_t + (\bar{s}_t + \theta\hat{s}_{t-1})^- / \lambda^i \\
&\quad + \lambda^e(\bar{s}_t + \theta\hat{s}_{t-1})^+, -\kappa^i \}) \\
&\quad + \eta_t^e (\min \{ S - \alpha s_t - (\bar{s}_t + \theta\hat{s}_{t-1})^- / \lambda^i \\
&\quad - \lambda^e(\bar{s}_t + \theta\hat{s}_{t-1})^+, \kappa^e \} / \lambda^e) \\
&\quad + E [J_{t+1}^4(\epsilon_t, \alpha s_t + (\bar{s}_t + \theta\hat{s}_{t-1})^- / \lambda^i \\
&\quad + \lambda^e(\bar{s}_t + \theta\hat{s}_{t-1})^+ + \text{sat}(\epsilon_t^- / \lambda^i + \lambda^e \epsilon_t^+, \alpha s_t \\
&\quad + (\bar{s}_t + \theta\hat{s}_{t-1})^- / \lambda^i + \lambda^e(\bar{s}_t + \theta\hat{s}_{t-1})^+))]) \\
\text{s. t.} \quad &0 \leq \alpha s_t + (\bar{s}_t + \theta\hat{s}_{t-1})^- / \lambda^i \\
&+ \lambda^e(\bar{s}_t + \theta\hat{s}_{t-1})^+ \leq S \\
&\lambda^i \kappa^i \leq \bar{s}_t + \theta\hat{s}_{t-1} \leq \kappa^e / \lambda^e
\end{aligned}$$

As in Section 4.4.1.3.2, we apply a similar substitution, but now with the term  $\theta\hat{s}_{t-1}$  included:  $y^i = \alpha s_t + (\bar{s}_t + \theta\hat{s}_{t-1})^- / \lambda^i$  and  $y^e = \alpha s_t + \lambda^e(\bar{s}_t + \theta\hat{s}_{t-1})^+$ . As before, the policy will be obtained by applying the reverse substitution, except that now the resulting control may be infeasible. We correct this possibility by truncating the resulting set-points at capacity limits:

$$\begin{aligned}
\tilde{z}_t^i(s_t) &= \text{sat}(z_t^i(0, s_t) - \alpha s_t - \theta\hat{s}_{t-1} / \lambda^i, \alpha s_t) \\
\tilde{z}_t^e(s_t) &= \text{sat}(z_t^e(0, s_t) - \alpha s_t - \lambda^i \theta\hat{s}_{t-1}, \alpha s_t)
\end{aligned}$$

Proceeding as in Section 4.4.1.3.2, we have the following suboptimal policy:

#### Policy 4 (Small arbitrage)

$$\begin{aligned}
\mu_t^4(\hat{s}_{t-1}, s_t) &= \\
&\begin{cases} \lambda^i \tilde{z}_t^i(s_t) \text{ if } c_t \alpha (\lambda^i - 1 / \lambda^e) s_t \\ \geq G_t^e(0, z_t^e(0, s_t)) - G_t^i(0, z_t^i(\hat{s}_{t-1}, s_t)) \\ \tilde{z}_t^e(s_t) / \lambda^e \text{ if } c_t \alpha (\lambda^i - 1 / \lambda^e) s_t \\ \leq G_t^e(0, z_t^e(0, s_t)) - G_t^i(0, z_t^i(0, s_t)). \end{cases}
\end{aligned}$$

Like **Policy 3**, this suboptimal policy is affine in  $\hat{s}_{t-1}$ . Note that this formulation does not neglect arbitrage, but rather gives priority to absorbing variability. As with **Policy 3**, one could apply the approximations of **Policy 2** to obtain a policy affine in both  $s_t$  and  $\hat{s}_{t-1}$ .

We now comment briefly on price uncertainty. If storage is small enough to be a price taker and  $c_t$  is uncorrelated through time, we may simply replace  $c_t$  by its expectation and not introduce any additional complexity to the model. Correlations may be incorporated by appending new price states; in this case, the new states must be discretized to obtain the optimal control, but the above suboptimal policies will remain respectively affine in  $s_t$  and  $\hat{s}_t$ . If storage is large enough to exert influence on the market price so that it is a function  $c_t(\bar{s}_t)$ , the affine structure will be lost due to the resulting products  $c_t(\bar{s}_t) s_t$ .

**4.4.1.3.4 Hybrid storage** Hybrid storage configurations have been proposed, in which different technologies with complementary strengths may achieve better performance than would any individual technology [60]; for example, high-energy, low-power capacity compressed air and low-energy, high-power capacity supercapacitors has been identified as a promising combination [40]. We now consider an extension to the above formulation in which a single facility chooses from a set of storage types at each time step, each with their own set of efficiency and capacity characteristics from Table 2. The formulation is an idealization in that by modeling the composite system as having a single energy state, we implicitly assume each unit in a hybrid configuration can arbitrarily exchange energy without losses.

Let  $Q$  be the set of available storage types. Also associated with each option are cost-of-use functions,  $v_q^i(x)$  and  $v_q^e(x)$  for  $q \in Q$ , which are additively separable (i.e.  $v_q^i(x_1 + x_2) = v_q^i(x_1) + v_q^i(x_2)$ ); these represent maintenance and lifetime costs for each technology. We can then redefine (4) and (5)

$$\begin{aligned} G_t^{i,q}(\hat{s}_{t-1}, y^i) &= \bar{g}_t(\hat{s}_{t-1}, y^i) - c_t \lambda_q^i y^i - v_q^i(\lambda_q^i y^i) \\ &\quad + \eta_t^i (\lambda_q^i \min \{y^i, -\kappa_q^i\}) \\ &\quad + E [J_{t+1}(y^i + \text{sat}((\theta \hat{s}_{t-1} + \epsilon_t)^- / \lambda_q^i \\ &\quad + \lambda_q^e (\theta \hat{s}_{t-1} + \epsilon_t)^+, y^i))] \\ G_t^{e,q}(\hat{s}_{t-1}, y^e) &= \bar{g}_t(\hat{s}_{t-1}, y^e) - c_t y^e / \lambda_q^e - v_q^e(y^e / \lambda_q^e) \\ &\quad + \eta_t^e (\min \{S - y^e, \kappa_q^e\} / \lambda_q^e) \\ &\quad + E [J_{t+1}(y^e + \text{sat}((\theta \hat{s}_{t-1} + \epsilon_t)^- / \lambda_q^i \\ &\quad + \lambda_q^e (\theta \hat{s}_{t-1} + \epsilon_t)^+, y^e))] , \end{aligned}$$

and rewrite (6) as

$$\begin{aligned} J_t^H(\hat{s}_{t-1}, s_t) &= \max_{q \in Q} \max \left\{ v_q^i(\lambda_q^i \alpha s_t) + c_t \lambda_q^i \alpha s_t \right. \\ &\quad + \max_{y^i \in \beta_q^i(s_t)} G_t^{i,q}(\hat{s}_{t-1}, y^i), \\ &\quad v_q^e(\alpha s_t / \lambda_q^e) + c_t \alpha s_t / \lambda_q^e \\ &\quad \left. + \max_{y^e \in \beta_q^e(s_t)} G_t^{e,q}(\hat{s}_{t-1}, y^e) \right\} . \end{aligned}$$

Some energy storage types, for example lead-acid batteries, are degraded by frequent cycling (repeated charging and discharging) [35]. Constant functions which index the number of cycles may therefore also be appropriate choices for  $v_q^i(x)$  and  $v_q^e(x)$ , in which case the function will only appear in  $J_t(s_t)$  and not  $G_t^{i,q}$  or  $G_t^{e,q}$  after the substitution. Note the same pair of substitutions from Section 4.4.1.3.2 has been applied, and via an identical construction, it can be shown that the optimal policy is of the form given in **Policy 1**, but with slightly more conditional statements.

**4.4.1.3.5 Single period approximation** We now combine approximations to obtain a model of storage behavior simple enough to combine with other models. We now assume that energy imbalances are uncorrelated and that there are no capacity premiums, i.e.  $\hat{s}_t = \epsilon_t$  and  $\eta^e(x) = \eta^i(x) = 0$ . In certainty equivalent control, disturbances are replaced with nominal values such as the mean or mode in the state evolution [12]; we apply this concept by replacing sat functions with their expectations, denoted

$$h_t(y) = E [\text{sat}(\hat{s}_t, y)] ,$$

and are left with a deterministic dynamic programming problem. We remark that simply replacing  $\hat{s}$  with zero, eliminating the sat function terms entirely, is inadequate: policies will be biased towards maintaining full capacity, because energy will appear to be sold without reducing holdings. Suppose that  $\lambda^i = \lambda^e = 1$ , and now also that  $\kappa^i = -\infty$  and  $\kappa^e = \infty$ , which is to say that there are no limits of charge and discharge rates. The certainty equivalent, perfectly efficient version of (6) is given by

$$J_t^{CEC}(s_t) = c_t \alpha s_t + \max_{0 \leq y \leq S} \{ -c_t y + \bar{g}_t(y) + J_{t+1}^{CEC}(y + h_t(y)) \},$$

which may be expanded into

$$J_t^{CEC}(s_t) = c_t \alpha s_t + \max_{\substack{0 \leq y_k \leq S \\ k = t, \dots, T-1}} \left\{ \bar{g}_T(y_{T-1} + h_{T-1}(y_{T-1})) + \sum_{k=t}^{T-1} [\bar{g}_k(y_k) + c_{k+1} \alpha h_k(y_k) + (c_{k+1} \alpha - c_k) y_k] \right\}$$

This is a separable optimization problem, and hence the associated cost and control,  $J_t^{CEC}$  and  $\mu_t^{CEC}$ , may be obtained by solving  $T - t$  nonlinear programming problems.

We now consider the infinite horizon analog, which, in addition to informing planning problems, admits an analytically tractable policy that we make use of in the next section. Let  $T = \infty$ ,  $c_t = c$ ,  $\bar{g}_t = \bar{g}$ ,  $h_t = h$ , and introduce the discount factor  $\gamma \in [0, 1)$ ;  $\gamma$  might reflect a variety of factors, such as component maintenance and investment of future profits. Because parameters no longer vary through time, each stage is identical. Let

$$z^\infty = \operatorname{argmax}_{0 \leq y \leq S} \bar{g}(y) + c \alpha h(y) + c(\alpha - 1)y. \quad (10)$$

The first two terms represent the expected profits from absorbing deviations, without consideration of arbitrage, and the last term is the loss due to leakage. The optimal cost is given by

$$\begin{aligned} J^{CEC}(s) &= c \alpha s + \sum_{k=0}^{\infty} \gamma^k (\bar{g}(z^\infty) + c \alpha h(z^\infty) + c(\alpha - 1)z^\infty) \\ &= c \alpha s + \frac{\bar{g}(z^\infty) + c \alpha h(z^\infty) + c(\alpha - 1)z^\infty}{1 - \gamma}. \end{aligned}$$

Assuming that  $\gamma$  is close to one,  $J^{CEC}(0)$  is a rough estimate for the lifetime profit gained from absorbing variability with storage. Since the  $J^{CEC}(0)$  is highly sensitive to  $\gamma$ , an ad-hoc parameter, the result may be more useful as a relative comparison between options. An additional convenience of this formulation is that  $\bar{g}(y) + c \alpha h(y)$  is concave if  $\alpha \geq 1 - \delta_s^e/c$ , which can be seen from the second derivatives. The dynamic behavior of a storage facility has been reduced to a single variable, convex optimization problem.

**4.4.1.4 Equilibria between storage and a producer** We now use the single period approximation of the previous section to examine the interplay between a variable generator and any number of storage units by identifying Nash equilibria [43]. Thus far, we have made no assumptions regarding the probability distribution of  $\hat{s}$ , nor have we discussed production scheduling,  $\bar{w}$ . Here, rather than defining a distribution on  $\hat{s}$ , we more realistically assume a distribution on net production,  $f(w)$  (and cumulative distribution  $F(w)$ ), e.g. a Weibull-based distribution for a wind power producer [45]. Since we are now considering a single time step, we will drop all time subscripts  $t$ .

In applying game theory, we implicitly assume that the producer and each storage is aware of each others' physical characteristics and state of charge; while the former is a reasonable assumption given that most storage technologies are well-characterized, the latter is an idealization necessary for the application of game theory. A potential future direction to add realism is modeling state of charge and other unknown parameters using Bayesian Nash equilibrium concepts [33].

The producer seeks to maximize its profits, as given by the expectation of (1), and the storage employs the infinite horizon strategy given by (10) in Section 4.4.1.3.5, under the assumptions that injection and extraction are perfectly efficient and that there are no limits on charging or discharging, and the earlier assumption that energy imbalances are uncorrelated. We assume that each renewable producer follows the optimal newsvendor contract strategy of [15, 48]. There is now a ladder of fees for available storage facilities and the transmission system, and variability is hence allocated in order of increasing price to the storage resources and then to the more expensive transmission system (the allocation of Section 4.4.1.2.3).

Storage facilities are indexed in order of increasing (price) fees so that  $\delta_1^i < \dots < \delta_N^i$ , where  $N$  corresponds to the transmission system, and we have assumed for convenience that extraction fees have the same ordering. Production and storage expected utilities are denoted  $\bar{u}_w$  and  $\bar{u}_1, \dots, \bar{u}_{N-1}$ . It can be seen in the Appendix from looking at second derivatives that  $\bar{u}_j$  is concave if  $\alpha \geq 1 - \delta_j^e/c$ , which is to say that expected losses from leakage are outweighed by profits from absorbing excess; we assume that this is the case throughout this section. Since  $\bar{u}_w$  is also concave, the existence of a pure Nash equilibrium is guaranteed [23, 29, 32]. Ideally (for analysis), we would like any equilibria to be unique, so that the behavior it describes is the likely outcome rather than one of many. Unfortunately, it does not appear that a unique equilibrium may be expected in general, but we are able to give sufficient conditions in some special cases.

Because the utilities are concave, a Nash equilibrium  $(\bar{w}^*, \bar{s}^*)$  is characterized by

$$\begin{aligned} \bar{s}_j^* &= \operatorname{argmax}_{\bar{s}_j} \bar{u}_j(\bar{w}^*, \bar{s}_j, \bar{s}_{-j}^*) \quad \text{s. t.} \quad 0 \leq \bar{s}_j + \alpha_j s_j \leq S_j \\ \bar{w}^* &= \operatorname{argmax}_{\bar{w}} \bar{u}_w(\bar{w}, \bar{s}^*) \quad \text{s. t.} \quad \bar{w} \geq 0 \end{aligned}$$

**Theorem 1** Suppose  $0 < \bar{s}_j^* + \alpha_j s_j^* < S_j$  for each storage  $j = 1, \dots, N - 1$ . Then the equilibrium is unique.

**Proof** If  $0 < \bar{s}_j + \alpha_j s_j < S_j$ , the equilibrium conditions are given by

$$\begin{aligned} \frac{d\bar{u}_j}{d\bar{s}_j} &= 0 \quad j = 1, \dots, N - 1 \\ \frac{d\bar{u}_w}{d\bar{w}} &= 0 \quad \bar{w} \geq 0. \end{aligned}$$

Now apply the substitution  $z_j = \bar{w} - \sum_{k=1}^j \bar{s}_k + \alpha_k s_{tk}$ ; note that each  $\bar{u}_j$  is also concave in  $z_j$ . Let  $z_j^*$  be

the unique solution to

$$\begin{aligned}
\frac{d\bar{u}_j}{dz_j} &= -\frac{d\bar{u}_j}{ds_j} \\
&= -(\delta_j^i - (\alpha_j - 1)c) F(z_j) \\
&\quad + (\delta_j^e + (\alpha_j - 1)c) \left(1 - F\left(z_j + \sum_{k=1}^j S_k\right)\right) \\
&\quad - c(\alpha - 1) \\
&= 0.
\end{aligned}$$

Since the constraints are not active,  $z_j^* = \bar{w}^* - \sum_{k=1}^j \bar{s}_k^* + \alpha_k s_k$ , where  $\bar{w}^*$  is maximum of zero and the solution to

$$\begin{aligned}
\frac{d\bar{u}_w}{d\bar{w}} &= \sum_{j=1}^{N-1} \left[ (\delta_j^i - \delta_{j+1}^i) F(z_j^*) \right. \\
&\quad \left. + (\delta_j^e - \delta_{j+1}^e) F\left(z_j^* + \sum_{k=1}^j S_k\right) \right] \\
&\quad + \delta_N^e - (\delta_1^i + \delta_1^e) F(\bar{w}) \\
&= 0.
\end{aligned}$$

We may then iteratively back-substitute for the equilibrium storage behavior via

$\bar{s}_j^* = \bar{w}^* - z_j^* - \sum_{k=1}^{j-1} \bar{s}_k^* + \alpha_k s_k$ . Since each  $z_j^*$  and therefore  $\bar{w}^*$  and  $\bar{s}_j^*$  are unique, the associated equilibrium  $(\bar{w}^*, \bar{s}^*)$  is unique.

The above argument describes a procedure for finding a candidate equilibrium, which is verified if no constraints are violated. The substitution is similar to that in Section 4.4.1.3.2, and its resultant decoupling is what guarantees uniqueness. The condition that no constraints be active may appear artificial, but in fact is indicative of how well designed the system is. For simplicity, suppose  $\alpha = 1$  and  $\delta_j^i = \delta_j^e$  for all  $j$ . No constraints are active at the equilibrium if the derivative at the lower constraint is positive and at the upper negative; more precisely, we can state this as

$$\begin{aligned}
\frac{d\bar{u}_j}{d\bar{s}_j} (-\alpha_j s_j) &= F\left(\bar{w}^* - \sum_{k=1}^{j-1} \bar{s}_k^* + \alpha_k s_k\right) \\
&\quad + F\left(\bar{w}^* + S_j + \sum_{k=1}^{j-1} S_k - \bar{s}_k^* - \alpha_k s_k\right) \\
&\geq 1
\end{aligned}$$

and

$$\begin{aligned}
\frac{d\bar{u}_j}{d\bar{s}_j} (S_j - \alpha_j s_j) &= F\left(\bar{w}^* - S_j - \sum_{k=1}^{j-1} \bar{s}_k^* + \alpha_k s_k\right) \\
&\quad + F\left(\bar{w}^* + \sum_{k=1}^{j-1} S_k - \bar{s}_k^* - \alpha_k s_k\right) \\
&\leq 1
\end{aligned}$$

The first condition says that when storage  $j$  is empty, the probability of being asked for reserves is greater than that of not having enough capacity - so it should increase its holdings. Similarly, the second condition says that when full, the probability of running out of reserves is less than that of being asked for capacity. As  $S_j$  is decreased, the first condition weakens, and as it is increased, then does the second; in other words, there is a preferred range of sizes at which storage is best prepared for deviations, and at which the equilibrium is unique.

When only one storage option exists, the situation simplifies so that uniqueness does not depend on the constraints.

**Theorem 2** *Suppose there is only one storage, and that  $\alpha_1 = 1$ ,  $\delta_1^i = \delta_1^e = \delta_1$  and  $\delta_2^i = \delta_2^e = \delta_2$ . Then the equilibrium is unique.*

**Proof** We show uniqueness using the sufficient condition from [52], which is as follows. Let  $\nabla$  be the gradient operator with respect to  $(\bar{s}_1, \dots, \bar{s}_{N-1}, \bar{w})$ , and define

$$G(\bar{w}, \bar{s}, r) = \begin{bmatrix} r_1 \nabla \frac{d\bar{u}_1}{d\bar{s}_1} \\ \vdots \\ r_{N-1} \nabla \frac{d\bar{u}_{N-1}}{d\bar{s}_{N-1}} \\ r_N \nabla \frac{d\bar{u}_w}{d\bar{w}} \end{bmatrix}.$$

If there exists an  $r \in \mathbb{R}^N$ ,  $r > 0$  for which the matrix  $G(\bar{w}, \bar{s}, r) + G(\bar{w}, \bar{s}, r)^T$  is negative definite, the equilibrium is unique.

In the present case,  $G(\bar{w}, \bar{s}, r)$  is given by

$$\begin{aligned} G(\bar{w}, \bar{s}, r) &= \begin{bmatrix} -r_1 \delta_1 & r_1 \delta_1 \\ r_2 (\delta_2 - \delta_1) & r_2 (\delta_1 - \delta_2) \end{bmatrix} \\ &\quad \times (f(\bar{w} - \bar{s} - s) + f(\bar{w} + S - \bar{s} - s)) \\ &\quad - \begin{bmatrix} 0 & 0 \\ 0 & 2r_2 \delta_1 \end{bmatrix} f(\bar{w}). \end{aligned}$$

Choose  $r_1 = \delta_2/\delta_1 - 1$  and  $r_2 = 1$ . The eigenvalues of the  $G(\bar{w}, \bar{s}, r) + G(\bar{w}, \bar{s}, r)^T$  are then zero and  $4(\delta_1 - \delta_2)$ . Since  $f(\bar{w} - \bar{s} - s) + f(\bar{w} + S - \bar{s} - s)$  and  $f(\bar{w})$  are both positive and the second matrix has different eigenvectors than the first,  $G(\bar{w}, \bar{s}, r) + G(\bar{w}, \bar{s}, r)^T$  is negative definite, and the equilibrium is unique.

With uniqueness guaranteed, we are able to examine equilibrium behaviors, which we do numerically in the next section.

#### 4.4.1.5 Examples

**4.4.1.5.1 Suboptimal policies under Gaussian uncertainty** In this example, we compare the profits obtained by **Policies 1** through **4** in a simple scenario with correlated Gaussian uncertainty.

Storage parameters are  $S = 1$ ,  $\delta_t^e(x) = \delta_t^i(x) = x$ ,  $\eta_t^e(x) = \eta_t^i(x) = 0$ , and  $\alpha = 0.95$ . Prices are given by  $c_t = 1 - \sigma \cos(\pi t/3)$  where  $T = 11$ ,  $J_T = 0$ , and  $\sigma = 1$ . At each time step,  $\epsilon_t$  is sampled from a zero mean normal distribution of variance 0.1, and the autocorrelation is  $\theta = 1$ . In **Policy 1**, both states are

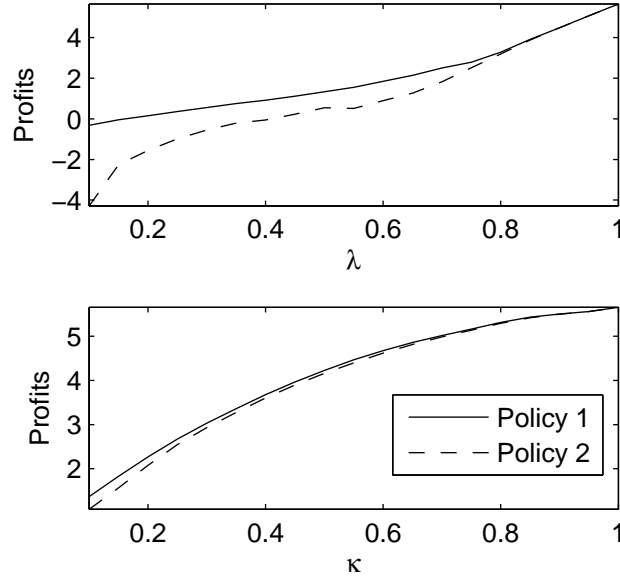


Figure 1: Performance of **Policy 2** relative to optimal as a function of inefficiency (top) and charge rate limits (bottom).

discretized to one hundred evenly spaced grid points. Under **Policy 2**,  $\hat{s}_t$  is discretized to the same resolution, but because  $s_t$  appears affinely, it does not need discretization. Expectations are evaluated numerically using Gauss-Hermite quadrature [34], and ten thousand Monte Carlo trials of each scenario were performed.

In Fig. 1, we examine the performance of **Policy 2** as a function of inefficiencies,  $\lambda^i = \lambda^e = \lambda$ , and charge rate limits  $\kappa^i = \kappa^e = \kappa$ ;  $\kappa = \infty$  and  $\lambda = 1$  while the other is varied in the top and bottom figures, respectively. We see that **Policy 2** is a good approximation when  $\lambda > 0.75$ , and that it remains a good approximation even for strict charge rate limits.

We now look at how suboptimal **Policies 3** and **4** become as the autocorrelation coefficient  $\theta$  (top) and price volatility  $\sigma$  (bottom) increase from zero to one. The above parameters are used, with efficiencies and charge rates fixed at  $\lambda_t^i = \lambda_t^e = 0.8$  and  $\kappa^i = \kappa^e = 0.75$ , and with  $\sigma = 1$  in the top plot and  $\theta = 0.5$  in the bottom plot. In **Policies 3** and **4**, only the state  $s_t$  is discretized, again to one hundred grid points.

We see in Fig. 2 that when correlations are large or when price volatility is small, **Policy 4**, which anticipates correlations, outperforms **Policy 3**, which neglects them.

**4.4.1.5.2 Efficiency of equilibrium with a wind energy producer** An important comparison to be made in competitive environments is the difference in social welfare, in this case the sum of all players' profits, when decisions are made centrally via optimization and competitively as the result of Nash equilibrium; such measures are sometimes called “the price of anarchy” [53].

In this example, we examine the equilibrium inefficiency between a wind energy producer and storage. Because the infinite horizon approximation is used in equilibrium analysis, both storage and the wind producer are effectively using single stage strategies, and the wind power forecast  $f(w)$  corresponds only



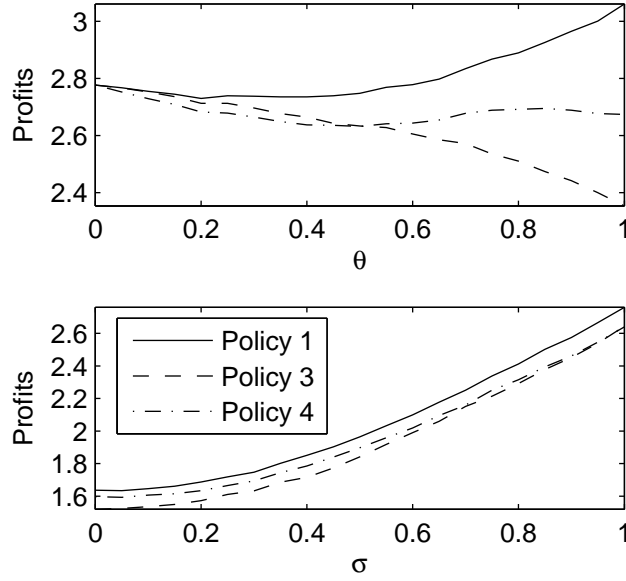


Figure 2: Performance of **Policies 3 and 4** relative to optimal as a function of correlations (top) and market price volatility (bottom).

to the particular time; given the high degree of approximation, our goal is to gain qualitative insight on the interaction between a producer and storage. Specifically, we investigate the efficiency of the Nash equilibrium discussed in Theorem 1 as a function of leakage,  $\alpha$ . Social welfare is defined to be  $\bar{u}_w + \bar{u}_1$ .

We model the distribution and speed versus power relationship following Ch. 4 of [45]; note that this is highly simplified in that we are modeling power. Wind speed is commonly modeled with a Weibull distribution, given by

$$\text{prob}(v) = \left(\frac{l}{q}\right) \left(\frac{v}{q}\right)^{l-1} e^{-\left(\frac{v}{q}\right)^l},$$

where  $l$  is a “shape” parameter and  $q$  a “scale” parameter. Wind power is approximately given by  $\rho A v^3 / 4$ , where  $\rho$  and  $A$  respectively are air density and rotor swept area. The factor of 4 is an approximation implying 50% turbine efficiency, which is better than modern installations but below Betz’ theoretical limit of 59%, and that efficiency is constant over windspeed, which it generally isn’t. The shape parameter is most commonly two, and  $q$  is proportional to mean wind speed for fixed  $l$ . In this example, we have ten wind turbines, each with parameters  $A = \pi 100^2 / 4 \text{ m}^2$ ,  $\rho = 1.225 \text{ kg/m}^3$ , and  $q = 7$  (when not being varied). We assume a generic storage with capacity  $S = 1 \text{ MWh}$ , current holding  $s_t = 0.1 \text{ MWh}$ , and that prices are  $c = 1$ ,  $\delta_s^i = \delta_s^e = 0.4$ ,  $\delta_T^i = \delta_T^e = 0.6$ .

Using the procedure described in Theorem 1 of Section 4.4.1.4, we numerically evaluate the equilibrium quantities  $z^*$ ,  $\bar{w}^*$ , and  $\bar{s}^*$ , and then the maximum attainable social welfare as functions of storage leakage,  $\alpha$ . Fig. 3 shows the equilibrium quantities (top) and the social welfares resulting from central and Nash strategies (bottom).

As  $\alpha$  is increased, scheduled storage holdings increase because the risk of losing energy to leakage before it can be sold is decreased. The producer then also schedules more, knowing that storage will be able to

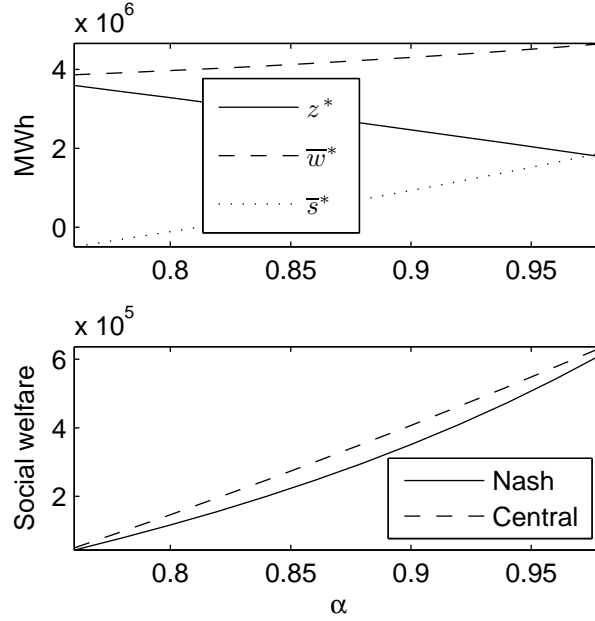


Figure 3: Resulting Nash equilibrium strategies (top) and social and competitive profits (bottom) as functions of storage leakage  $\alpha$ .

cover a larger shortfall. This is reflected in both the competitive and centrally optimal strategies. However, the centrally optimal strategy always results in greater social welfare than the Nash equilibrium. The difference is less severe when leakage is small and when  $\alpha = 3/4$ , below which it is optimal to empty storage and Theorem 1 no longer applies.

The reason the socially optimal and competitive strategies coincide below  $\alpha = 3/4$  is because once it is optimal for storage to empty its holdings, there is essentially only one decision maker. However, the results for larger  $\alpha$  indicate that as leakage shrinks, the objectives of storage and the producer become more compatible. Broadly interpreting this result, one may infer that physical inefficiencies of storage, which only directly affect the profits of storage, lead to inefficient market outcomes when other physical constraints are not in play; in other words, equilibria can be nearly efficient either when storage is physically efficient, or when it is so inefficient that its total contribution is very small.

**4.4.1.5.3 Optimal scheduling with a photovoltaic producer: absorbing deviations versus price arbitrage** Intemporal arbitrage over fluctuating market prices is built into the formulation of Section 4.4.1.3, and theoretically can exert a strong influence on the optimal policy. In this example, we examine the balance between absorbing deviations and arbitrage for the case of a single storage facility absorbing variability from a photovoltaic source.

Irradiance has been observed to follow bimodal distributions due to clouds [58], which we approximate with a Bernoulli random variable. Specifically, energy production at time  $t$  is given by  $w = b_t v_t$ , where  $v_t$  is a deterministic function and  $b_t$  a random variable equal to one with probability 0.6 and zero with probability 0.4. We assume for simplicity that the producer's forward contract is known to be its expected production,  $\bar{w}_t = pv_t$ , and thus  $\hat{s}_t = \hat{w}_t = w - pv_t$ .

We examine the optimal policy over a twelve hour period for a normalized system, over which the peak

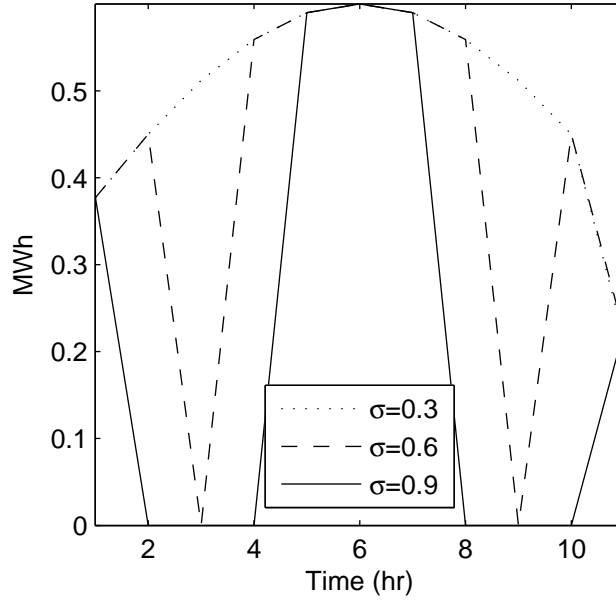


Figure 4: The optimal  $z_t$  for storage both absorbing deviations from solar and arbitraging over market prices as a function of arbitrage opportunity,  $\sigma$ .

solar output as  $v(t) = (1 + \sin(\pi t/12))/2$ . Storage capacity, leakage, and fee are  $S = 1$ ,  $\alpha = 0.9$ , and  $\delta = 0.1$ , and it is desirable to be at half-capacity at the end of the time period, so that

$$J_{12}(s) = 10(S/2 - s)^2. \quad (11)$$

The market price is  $c_t = 1 - \sigma \cos(\pi t/3)$ , where the parameter  $\sigma$  represents the level of arbitrage opportunity. We again assume that  $\eta_t^e(x) = \eta_t^i(x) = 0$ . Since correlations, inefficiencies, and charge rate constraints are not factors in this example, **Policies 1** through **4** are identical. We are interested in  $z_t$ , the optimal holding level determined by the optimal policy.

Fig. 4 shows the optimal trajectory of  $z_t$  for three values of  $\sigma$ : we see that as  $\sigma$  is increased, the market price fluctuations dominate storage behavior. This result suggests that if storage is to absorb deviations, fees must be allowed to be large enough to induce storage to pay less attention to arbitrage opportunities. We remark that this may indeed be the case; the uniform auction, in which the highest bid is paid to all, is a current common format for reserve procurement [28].

**4.4.1.6 Summary and future work** Heavy reliance on generator reserves diminishes the value of distributed generation. Local energy storage is a promising mechanism for absorbing the variability of renewable energy sources, a primary obstacle to integration of distributed generation. Since storage may enter electricity markets as independent, profit-seeking entities, it is necessary that financial structures exist through which storage and variable generators can efficiently interact.

We have analyzed a simple scenario in which independent storage either cooperates with an intermittent energy producer or competes in reserve markets. Our results are summarized as follows.

- The globally optimal ex-ante storage scheduling strategy is an inventory control-like policy, to which

minor approximations yield affine suboptimal policies.

- Under certain conditions, the approximate equilibrium between storage and production inherits an inventory control-like structure and is unique, allowing meaningful numerical examination. Results show that inefficiencies are a potential source of competitive behavior that does not align with socially optimal operation.

Extensions to the basic framework considered here include simultaneous fee and schedule determination and network considerations. More generally, a large number of distributed energy storage resources will soon be either directly or indirectly (economically) available to grid operators. Utilizing these resources to balance variability and increase system performance and robustness represents a rich new class of problems in power systems.

**4.4.1.7 Appendix: Equilibria with heterogeneous storage fees** In this appendix we write the derivatives of the expected production and storage utilities from Section 4.4.1.4. Storage attempts to maximize its infinite horizon stage profits as in Section 4.4.1.3.5, and the producer its expected profits. Production utility,  $\bar{u}_w$ , is then simply the expectation of (1), and the utility of storage  $j$ ,  $\bar{u}_j$ , is  $J^{CEC}$  of Section 4.4.1.3.5. The second derivatives show that each utility is concave; this of course because the pre-expectation utilities are concave functions of  $\bar{w}$  and  $\bar{s}_j$ , respectively, for any realization of the random variable  $w$ . The first and second derivatives (without the  $1 - \gamma$  discount term) are given by

$$\begin{aligned}
\frac{d\bar{u}_w}{d\bar{w}} &= \sum_{j=1}^{N-1} \left[ (\delta_j^i - \delta_{j+1}^i) F \left( \bar{w} - \sum_{k=1}^j \bar{s}_k + \alpha_k s_{tk} \right) \right. \\
&\quad \left. + (\delta_j^e - \delta_{j+1}^e) F \left( \bar{w} + \sum_{k=1}^j S_k - \bar{s}_k - \alpha_k s_{tk} \right) \right] \\
&\quad + \delta_N^e - (\delta_1^i + \delta_1^e) F(\bar{w}) \\
\frac{d^2 \bar{u}_w}{d\bar{w}^2} &= \sum_{j=1}^{N-1} \left[ (\delta_j^i - \delta_{j+1}^i) f \left( \bar{w} - \sum_{k=1}^j \bar{s}_k + \alpha_k s_{tk} \right) \right. \\
&\quad \left. + (\delta_j^e - \delta_{j+1}^e) f \left( \bar{w} + \sum_{k=1}^j S_k - \bar{s}_k - \alpha_k s_{tk} \right) \right] \\
&\quad - (\delta_1^i + \delta_1^e) f(\bar{w}) \\
\frac{d\bar{u}_j}{d\bar{s}_j} &= (\delta_j^i - (\alpha_j - 1)c) F \left( \bar{w} - \sum_{k=1}^j \bar{s}_k + \alpha_k s_{tk} \right) \\
&\quad - (\delta_j^e + (\alpha_j - 1)c) \\
&\quad \times \left( 1 - F \left( \bar{w} + \sum_{k=1}^j S_k - \bar{s}_k - \alpha_k s_{tk} \right) \right) \\
&\quad + c(\alpha - 1) \\
\frac{d^2 \bar{u}_j}{d\bar{s}_j^2} &= -(\delta_j^i - (\alpha_j - 1)c) f \left( \bar{w} - \sum_{k=1}^j \bar{s}_k + \alpha_k s_{tk} \right) \\
&\quad - (\delta_j^e + (\alpha_j - 1)c) \\
&\quad \times f \left( \bar{w} + \sum_{k=1}^j S_k - \bar{s}_k - \alpha_k s_{tk} \right).
\end{aligned}$$

## 4.4.2 Investigating the Value of Energy Storage in the Western Interconnection

**4.4.2.1 Introduction** The electric power system in the United States is facing a myriad of systemic changes that have the potential to disrupt current business operations. Renewable energy systems are increasingly being deployed across the country, and with the addition of their low-cost power comes a potential decrease in system reliability due to renewable energy variability and uncertainty. Energy storage could mitigate these problems. Depending on the type of storage technology, energy storage has the potential to decrease the need for conventional generators to serve peak load and deliver ancillary services, postpone infrastructure investment, provide ramp support for renewables, and increase system reliability [36]. At present, energy storage resources are limited, coming primarily from hydropower and having produced only 5.9% of the net electricity generated in 2008 [42]. Other technologies under consideration include compressed air energy storage (CAES), sodium sulfur batteries, fly wheels, and lithium ion batteries, but their portion of current storage capacity relatively small at the moment [36]. The buildout of these newer technologies will depend at least in part on their potential to provide value to the current electricity system.

The value of storage performing particular electricity system functions has been addressed in the literature. Sioshansi et. al. investigate the value of the energy arbitrage function of storage, and demonstrate that the value of energy arbitrage decreases as the total capacity of storage on the system is increased [57]. Eyer and Corey discuss 26 different value sources for storage, and benefit ranges in \$/kW for each source [27]. Drury, Denholm, and Sioshansi look at the revenue potential for CAES operators, if they participate in additional markets in addition to the energy market [26].

We present a model that assesses the value to the electricity system of storage additions that can perform several different electricity functions. Our model adds storage to the grid system, and uses the added resources to provide regulation, load following, and arbitrage. For each hour, the allocation of services to be provided by storage is determined endogenously, based on the most valuable action and the current operating constraints of the system. By allowing storage to provide both arbitrage and reserves, we are able to model simultaneous participation by storage in multiple markets, and can determine the value to the system of this behavior.

**4.4.2.2 Methods** The model devised for this analysis is an hourly unit commitment model of the Western Electricity Coordinating Council (WECC). The transmission network is modeled using DC load flow, and the optimization problem is formulated as a mixed-integer linear program, solved using a branch-and-cut algorithm that is implemented using the CPLEX 12.5 C++ library. The model objective, given in the Appendix, is to minimize operating costs subject to transmission and generator limits as well as reserve capacity requirements. The model is run independently for each of 366 days and produces, for each day, an hourly operation schedule that includes generator commitment and dispatch, reserve provision, power transfer along major transmission lines, and charging and discharging of storage devices, including existing pumped hydro plants.

The layout of the system network for the model is based on data for the 240-bus model created and published in association with the paper "Reduced Network Modeling in WECC as a Market-Based prototype", by James Price and John Goodin [50]. Additionally, the generation profiles for the output of hydro, biomass, wind, solar, and geothermal plants come from this model, as well as demand at each node in each hour. Figure 5 shows the yearly load duration curve that is modeled.

The disaggregated generator data from the Price model are also used in combination with generator

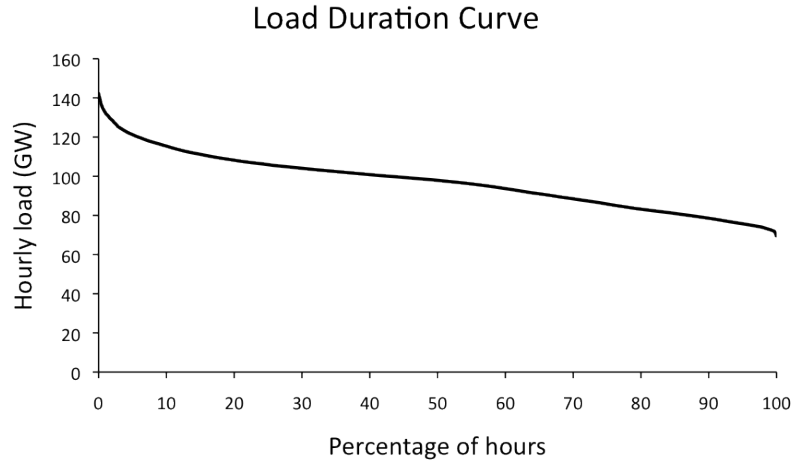


Figure 5: The load duration curve for the year of data modeled.

category and fuel data assumptions from WECC's Transmission Expansion Planning Policy Committee [22]. The data from the Price model are broken down into CAISO and Non-CAISO generators, for which different data are reported. For the Non-CAISO generators, a prime mover is identified, and can therefore be matched easily to the TEPPC data to obtain ramp limits, minimum up- and down-times, minimum operating capacities, start-up costs, and operations and maintenance costs. For the CAISO generators, only a fuel type is specified, and not a prime-mover. In these cases, we chose the generator type from the TEPPC data with the heat rate that was the closest to the heat rate reported in the Price model.

In total, our model commits 185 generators, of which 38 are coal-fired, 135 are gas-fired, 4 are nuclear, and 8 are run on fuel oil. The maximum regulation and load following capabilities of each generator were calculated based on the maximum generator movement in 10 minutes, using the one-minute ramp rate for the generator's prime mover. Generator limits on ramps between hours were calculated based on maximum generator movement in 60 mins [44].

In addition to thermal generators, the model also dispatches 4 pumped-hydro plants in all scenarios, as well as an added amount of extra storage that is between 0 GW/0GWh and 80GW/20 GWh. All added storage in this model is assumed to have an efficiency of 90% on both charge and discharge, to be able to fully charge and discharge in one hour, and to have a power to energy ratio of 4. The efficiencies and capabilities for the pumped hydro plants are taken from the Price model. We allow both types of storage to provide regulation and load-following, subject to constraints that require enough energy to be present in the battery (or energy capacity for charging in the case of down reserves) for provision of the committed regulation capacity for 15 minutes of load following for two hours [56].

Storage capacity can be used to serve multiple electricity system functions, which is a key aspect of storage that is built into this model. In the model, storage capacity can be used in both the energy and reserve markets. From an operational perspective, this means that when storage capacity is in any given charge state, one portion of its stored energy can be allocated to providing reserve, while another portion can be allocated to providing energy that is actually serving demand. Storage devices can serve as up-reserve by increasing their rate of discharging and down-reserve by increasing their rate of charging. This is achieved in the model by two constraints: the first constraining the sum of reserve and energy provided by a storage device in a given hour to be at most the total energy stored in the device at the beginning of the

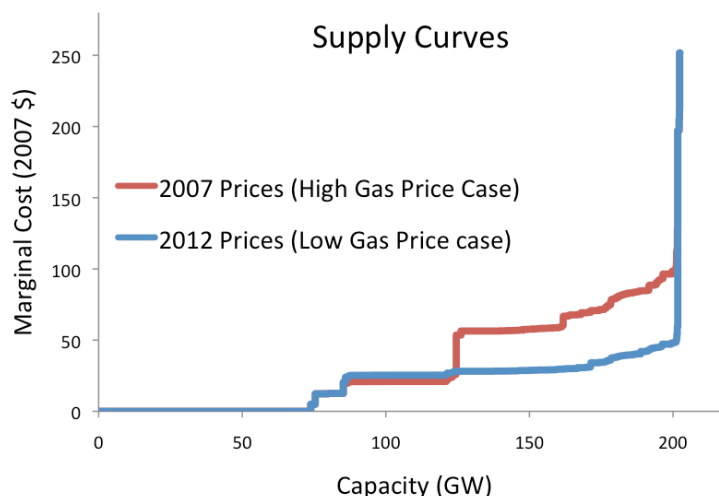


Figure 6: The supply curves for the set of generators modeled, in the low and high gas price scenarios. The supply curves in this graph are created for the Low Renewables scenarios; the High Renewables scenarios have a larger region where the marginal price is zero.

hour, as in Eq. (18), and the second constraining the total amount of down-regulation provided to be no more than the difference between the amount used to provide energy and the total available capacity in the storage, as in Eq. (21). Storage must have enough energy to provide whatever it commits to regulation for 15 minutes, and to load following for two hours, as in Eq. (19), (22), (20) and (23) [56].

Because it would otherwise be optimal to fully discharge storage devices at the end of the unit commitment modeling period, we first run a two-day unit commitment model in four hour increments to determine storage charge levels for the end of the first day, followed by a single-day unit commitment in one hour increments with fixed final storage charge levels and final generator operating states [56].

We determined the locations for newly added storage devices prior to running the unit-commitment model by slightly modifying the model and running it for the peak demand day. At each node, we added a decision variable denoting the total amount of energy storage capacity to be added at that node. We then constrained the sum of these decision variables to be less than or equal to the total amount of storage capacity being added for each storage increment. We performed this process iteratively, to simulate the incremental addition of storage devices to the grid. At each successive iteration, the next added amount of storage is allocated in the network such that it can optimally provide energy arbitrage and minimize congestion on the peak day.

We explored four different scenarios with this model, which allowed us to explore the effects on the value of storage of high vs. low natural gas prices and high vs. low penetrations of renewables. First, we explored low and high natural gas prices. For the low gas price scenario, we used \$3.17/MMBtu for gas and \$2.22/MMBtu for coal (both in 2007\$), which are average prices for these two fuels in 2012 [3]. For the high gas price scenario, we used \$7.12/MMBtu for gas and \$1.77/MMBtu for coal (both in 2007 \$), which are average prices for these fuels in 2007 [2]. Figure 6 shows supply curves for both the high and low gas price scenarios.

For each natural gas price, we also looked at a low-penetration renewables scenario and a high-penetration renewables scenario that meets California's renewable portfolio standard goals, as



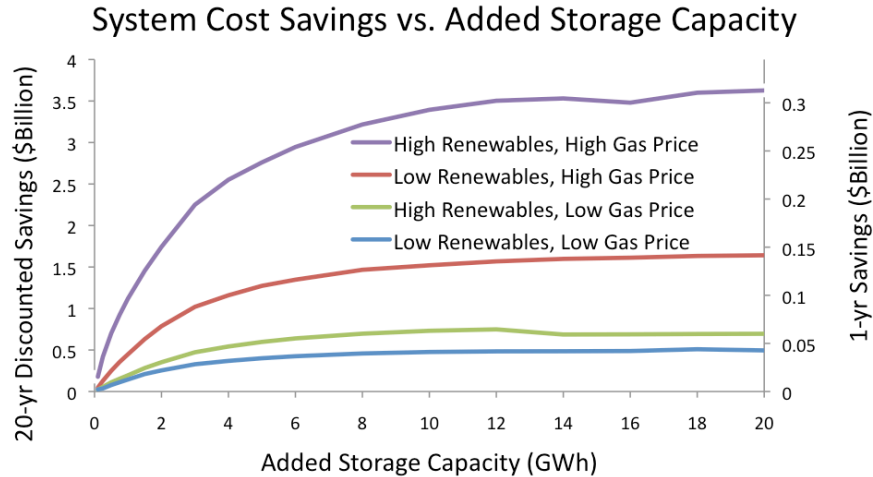


Figure 7: System cost savings as storage penetration is increased. System cost savings levels out in each scenario, and by the time 20 GWh of additional storage are added, increasing the amount of storage on the system no longer produces significant savings.

defined in [50]. The low penetration scenario has 6.5 GW of wind and 0.5 GW of solar, whereas the high-penetration scenario has 24 GW of wind, and 7 GW of solar.

#### 4.4.2.3 Results

**4.4.2.3.1 System Operating Cost Benefits** The total system operating costs for each scenario drop with increasing penetration of storage. For each scenario, total system cost savings are plotted in Figure 7. As one would expect, the largest savings are occur in the high gas / high renewables scenario and the smallest savings are in the low gas / low renewables scenario.

Using the total system costs for operating, we then calculate a marginal operating cost benefit for additional storage at each increment of additional storage (i.e. the system cost savings due to the added increment divided by the size of the increment). The changes are shown in Figure 8. These savings are originally calculated over the simulation year (right y-axis), but we also report a “20 year” savings, computed by assuming the same savings occur in each of 20 years, discounted to the present using a 7% discount rate. Clearly the infrastructure, demand and fuel prices will not be static over a twenty year horizon. But this metric gives us a sense of what the levelized operating cost benefits of the storage *could* be over a long operating period.

We note, however, that in these runs there is no supply scarcity – i.e. there is adequate capacity to meet demand and reserve requirements in all hours. Therefore for these runs there is no capacity value in additional amounts of storage. However future scenarios are likely to differ – for example if all California once-through-cooling plants retire or demand grows significantly at one or more points in the interconnection. We are in the process of evaluating such scenarios, but at this point the model cannot quantify the potential capacity value of storage in the future.

One might intuitively expect that the storage functions would always assigned to the highest value services first, which would imply a monotonically decreasing marginal benefit curve. Though the marginal

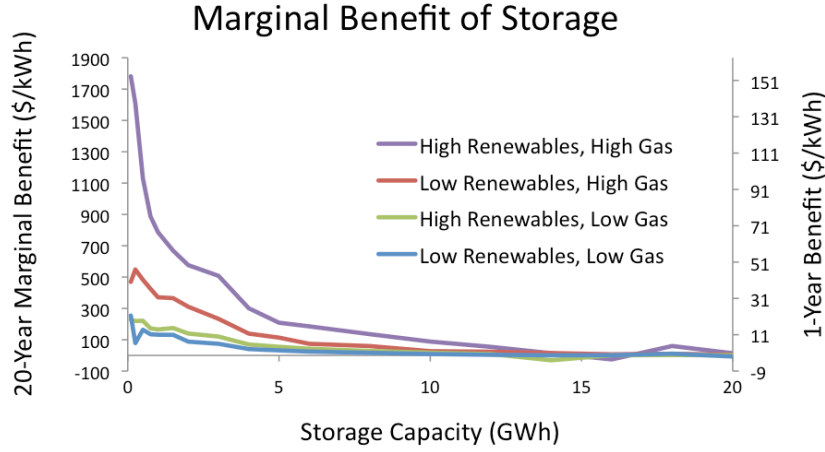


Figure 8: Marginal benefit of additional storage

benefit curve shows this general trend, it does have positive slope increments and even some slightly negative values. This characteristics can be attributed to the fact that the quality of the solution for each storage increment is variable.

**4.4.2.3.2 Implied Market Prices and Profits** To calculate the profits for the storage devices on the system, we first estimated the market clearing prices for each market for each hour of the day. For the energy market, each generator or storage device is paid the locational marginal price (LMP) for the node at which it is located. We obtain this price from the dual of the nodal balance constraint, Eq. (14), which we will call  $\lambda_{nt}$ , where  $n \in N$  and  $t \in T$ . Then, for each reserve market, including regulation up, regulation down, load following up, and load following down, the market clearing price in each hour is the maximum opportunity cost (\$/MW) faced by a generator that is providing the corresponding resource in that hour. We will refer to these hourly prices as  $\lambda_t^{ru}$ ,  $\lambda_t^{rd}$ ,  $\lambda_t^{lfu}$ , and  $\lambda_t^{lfd}$  for regulation up, regulation down, load following up, and load following down, respectively. Only generators constrained by their maximum capacities (for generators providing up reserves) or minimum capacities (for generators providing down reserves) experience opportunity costs. Generators that have not committed their full, currently available capacities are indifferent to committing their capacities to one market versus another; they have available capacity to do both. [62] We assume that all generators bid their opportunity costs into the reserve markets, and that all storage devices in these markets act as price takers.

The gross profit,  $Z_i$ , for a given storage device  $i$  over the entire year, then, is calculated as follows:

$$Z_i = \sum_{t \in T} \lambda_{it} e_{it}^{dr} - \lambda_{it} e_{it}^{cr} + \lambda_t^{ru} e_{it}^{ru} + \lambda_t^{rd} e_{it}^{rd} + \lambda_t^{lfu} e_{it}^{lfu} + \lambda_t^{lfd} e_{it}^{lfd} \quad (12)$$

The total gross profit in the system,  $Z$ , is the sum of the  $Z_i$ 's over all storage devices in the system. Gross profit is calculated as the revenue received in the energy, regulation, and load following markets, less the cost to charge storage using energy market prices.

Figure 9 shows the value of each component of  $Z$ , as additional storage devices are added to the system. The total revenue available is largest in the high renewables, high gas price case, when the reserve requirements are the largest due to the renewables, and the market clearing prices are set by generators

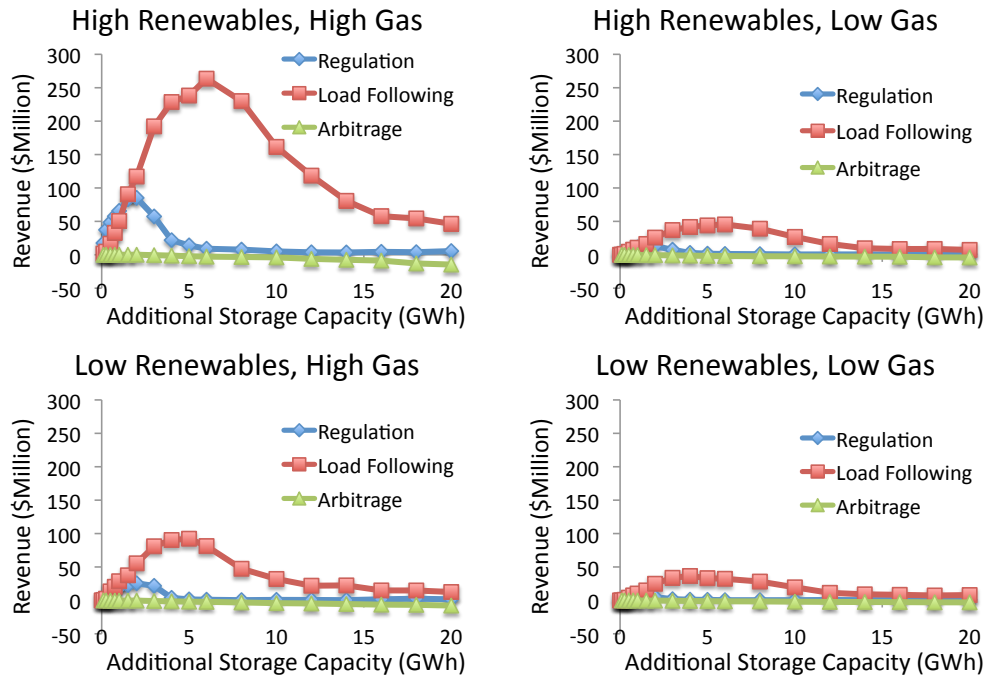


Figure 9: Contributions to gross profit from each service provided by storage. As storage is added to the system, the total revenue available for storage to obtain decreases.

with higher marginal fuel costs. The value to storage operators is coming from reserves more than arbitrage; in fact, as the total amount of storage on the system decreases, storage operators lose money on arbitrage in favor of participating in the regulation and load following markets.

Figure 10 shows profit per unit of energy storage as the total amount of storage in the system is increased. The total gross profit is largest in the high renewables, high gas price case, when the reserve requirements are the largest due to the renewables, and the market clearing prices are set by generators with higher marginal fuel costs. Figure 11 shows the proportion of various ancillary services requirements that is satisfied by storage in the High Renewables, High Gas Price case. The most valuable service, and the one storage satisfies first, is regulation up. The next most valuable are load following up and regulation down, and finally load following down. The Figure 12 shows the total energy transferred into and out of storage devices at each storage penetration. This gives an indication of the actual wear experienced by storage devices operated in this system, and shows that as storage penetration increases, storage is most utilized in the High Renewables, High Gas Price scenario.

There are several factors that allow storage to provide system cost savings as penetrations of storage are increased. In this model, these include arbitrage, reduction in costs due to fewer generator starts, and fewer generators being required to provide regulation because enough storage exists to provide such services. Figure 9 indicates that the arbitrage benefit is much smaller than the benefit to reserves provision. Figures 13 and 14 indicate that the number of generator starts and the cost due to generator starts both decrease as the amount of storage present in the system is increased.

As storage is added to the system, the carbon emissions associated with operating the system increase for

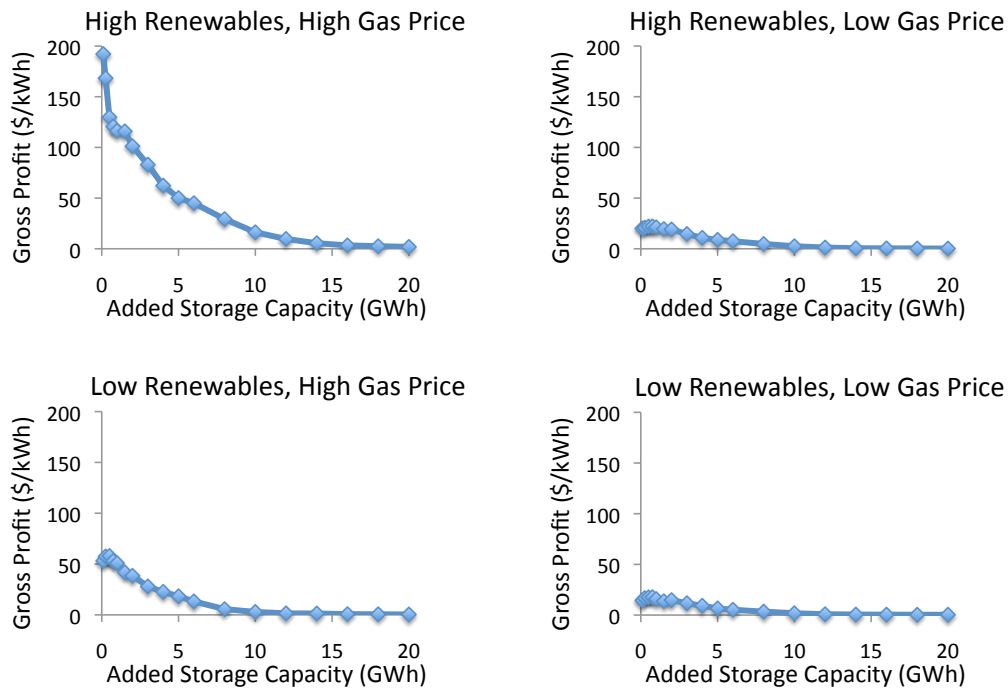


Figure 10: Gross profit per unit of energy storage. As storage is added to the system, the marginal gross profit seen by storage operators decreases.

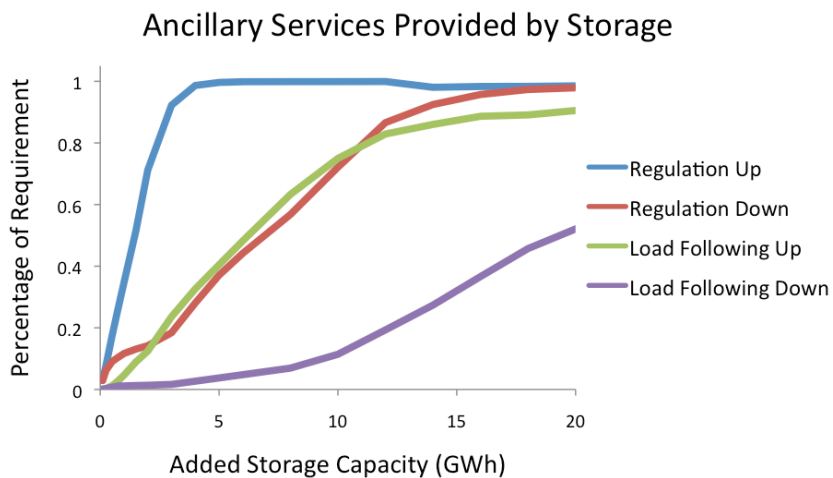


Figure 11: In the High Renewables, High Gas Price case, storage quickly moves to provide all required regulation up. Subsequently, storage emphasizes the provision of regulation down and load following up, and then finally begins to increase the proportion of load following down provided.

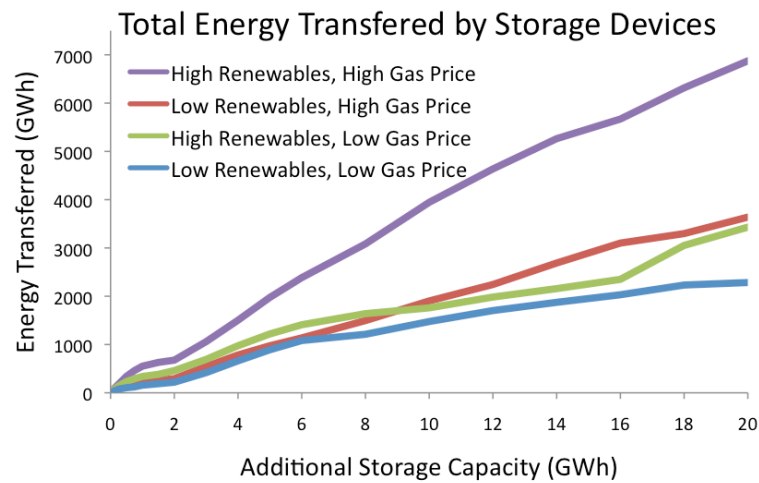


Figure 12: The total energy transferred is calculated as the sum of the energy charged to and discharged from the storage devices over the course of the year. The energy transferred to and from the storage devices grows fastest with storage penetration in the high renewables, high gas price case.

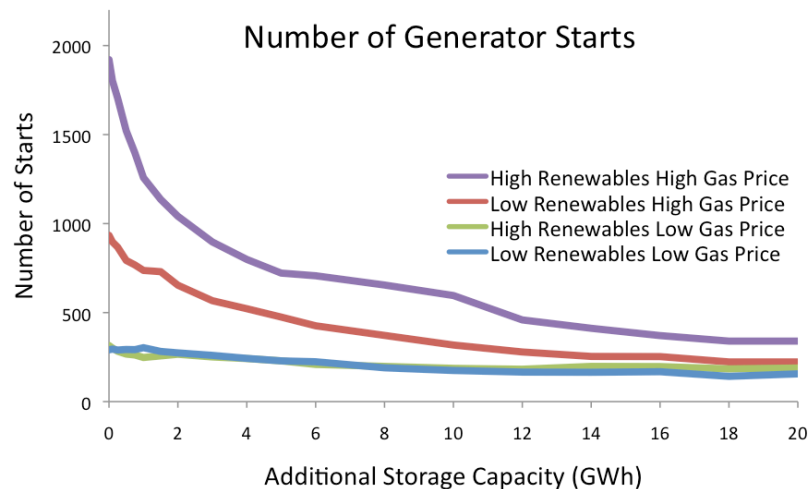


Figure 13: Number of generator starts per year. As the total amount of additional storage capacity on the system increases, the total number of generator starts decreases. The resulting reduction in startup costs paid contributes to the corresponding decreases in total system operating costs, as shown in Figure 7.

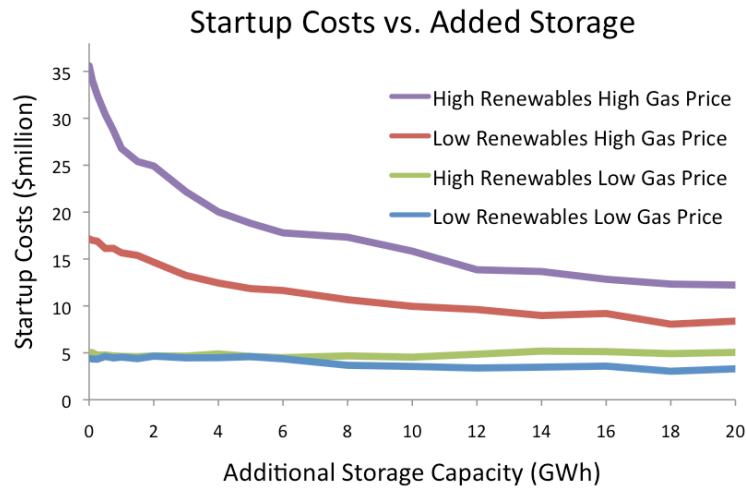


Figure 14: Cost of generator starts. As the total amount of additional storage capacity on the system increases, the total cost of generator starts decreases. The resulting reduction in startup costs paid contributes to the corresponding decreases in total system operating costs, as shown in Figure 7.

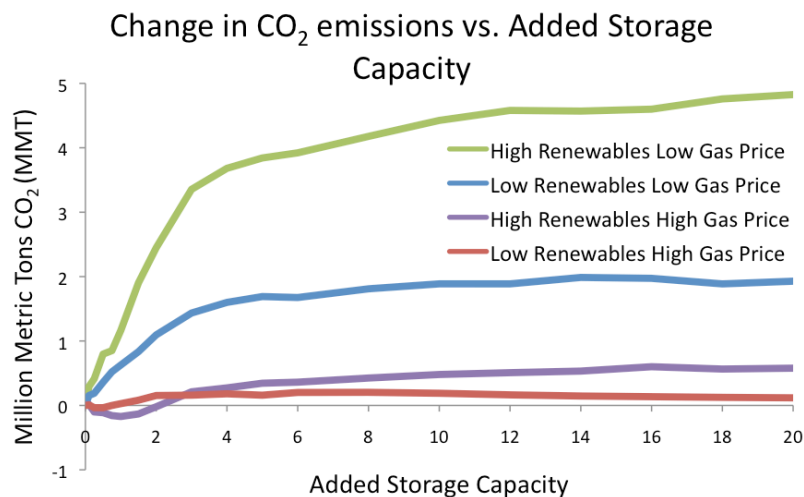


Figure 15: As storage is added to the system, the carbon dioxide released due to system operations increases. In 2005, WECC emissions were between 370 and 385 MMT [19].

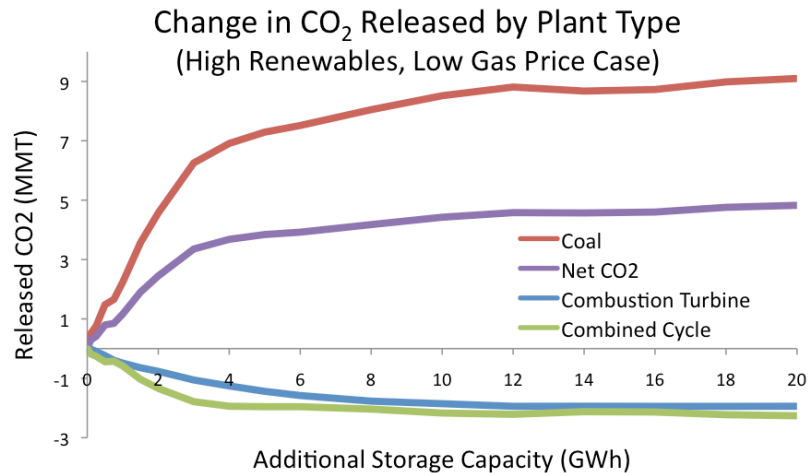


Figure 16: This figure shows the emissions sources for the scenario with the largest increase in CO<sub>2</sub> emissions from Figure 15, with high renewables penetration and a low gas price. As the amount of storage added is increased, emissions from coal plants rise and emissions from combined cycle and combustion turbine gas plants fall.

most scenarios. Figure 15 shows that carbon dioxide emissions strictly increase in all scenarios as storage penetration increases, with the exception of the High Renewables, High Gas Price scenario. In the High Renewables, High Gas Price scenario carbon dioxide emissions experience a slight decrease until 1GWh of storage capacity is present, and then emissions begin to increase. The scenario with the largest increase in emissions is the High Renewables, Low Gas Price scenario. In this scenario, coal plants experience an increase in operation, and combustion turbine and combined cycle gas plants experience a decrease in operation. As the fuel mix transitions towards coal, the overall carbon emissions of the system increase. This is shown in Figure 16.

**4.4.2.4 Concluding Remarks** With respect to system operation, storage additions have the most economic benefit in the high gas price scenarios, where there is a big discrepancy between the price of gas-fired peaker plants and coal-fired baseload plants (see the large jump in marginal costs for the 2007 fuel prices at 130 GW in Figure 6). When this is the case, the economic benefits of storage are larger, because storage is able to replace more expensive gas plants with energy from storage. Since the gas plants would operate as peakers or providing reserves due to their larger ramping capacity, these are the functions that storage takes over. Figure 9 shows that the most valuable functions for storage to take over are reserve functions.

Although high gas prices in combination with low coal prices is the first predictor of large economic benefits from storage, having high concentrations of renewables also predicts economic benefits for storage in this model. This is due to the increased reserve requirements for the system as a results of the larger share of intermittent resources present. Higher renewables do not have an effect due to prices, since higher renewables mean higher portions of generation with a zero marginal cost.

In combination, these factors indicate – unsurprisingly – that storage is most beneficial in a system that has both a large difference in marginal costs between low- and high-cost plants, and large reserve requirements. With the recent decrease in natural gas prices, the current system is moving away from an

advantageous price dichotomy, and closer to a system in which storage has a smaller economic effect. Current reserve requirements are in the process of being revised so that uncertainty in intermittent forecasts can be better accounted for, and as this process continues, it is likely that the system will experience an increase in value sources for storage.

Though the model results do not show evidence of large operating cost benefits per-kWh of storage capacity, we note that we have not quantified the potential capacity value of storage in future scenarios of demand scarcity. This is an important area of future work.

With respect to carbon emissions, the presence of storage on the system causes an increase in CO<sub>2</sub> emissions for all scenarios, except for at very small storage penetrations in the High Renewables, High Gas case. This is most likely due to increased usage of coal plants in lower demand, low price hours to charge storage devices. The cheapest times to charge storage devices will tend to be in hours when there are more coal plants on, which means that the energy stored in and then delivered by the storage devices will be dirtier than the energy supplied without storage. Relative to WECC-wide emissions, the increases are small; around a 1.4% increase in emissions from the 2005 level.

Overall, the benefits from increasing presence of storage quickly decline, regardless of the influence of gas prices or renewables penetration. By the time 10 GWh of storage exists on the system, operating cost benefits have largely ceased to be realized in all scenarios, and the carbon intensity of the system continues to increase as a result of the presence of storage. This suggests that any additions of storage should be smaller than 10 GWh, if economic benefits are a motivation for increased storage. We note, however, that we have not modeled the possible effect of California's carbon emissions cap, and this may have important implications on the model's carbon results.

## Appendix: Model Description

$$\min \sum_{g \in G} \sum_{t \in T} \gamma_g q_{gt} + SU_g s_{gt} \quad (13)$$

$$\text{s.t.} \quad \sum_{g \in G_n} (q_{gt}) + e_{nt}^{dr} - e_{nt}^{cr} + \sum_{i \in N} Y_{ni}(\theta_{nt} - \theta_{it}) = L_{nt} \quad \forall n \in N, t \in T \quad (14)$$

$$Y_{ij}(\theta_{it} - \theta_{jt}) \leq \bar{Z}_{ij} \quad \forall (i, j) \in (N, N), t \in T \quad (15)$$

$$0 \leq e_{nt}^l \leq \bar{E}_n \quad \forall n \in N, t \in T \quad (16)$$

$$e_{n,t+1}^l = e_{nt}^l + T_l \beta_n e_{nt}^{cr} - \frac{T_l}{\delta_n} e_{nt}^{dr} \quad \forall n \in N, t \in T \quad (17)$$

$$e_{nt}^l \geq \frac{T_l}{\delta_n} e_{nt}^{dr} + \frac{T_l}{\delta_n} e_{nt}^{ru} + \frac{T_l}{\delta_n} e_{nt}^{lfu} \quad \forall n \in N, t \in T \quad (18)$$

$$e_{nt}^l \geq \frac{15T_l}{60\delta_n} e_{nt}^{ru} + \frac{2T_l}{\delta_n} e_{nt}^{lfu} \quad \forall n \in N, t \in T \quad (19)$$

$$e_{n,t+1}^l \geq \frac{15T_l}{60\delta_n} e_{nt}^{ru} + \frac{2T_l}{\delta_n} e_{nt}^{lfu} \quad \forall n \in N, t \in T \quad (20)$$

$$\bar{E} - e_{nt}^l \geq T_l \beta_n e_{nt}^{cr} + T_l \beta_n e_{nt}^{rd} + T_l \beta_n e_{nt}^{lfd} \quad \forall n \in N, t \in T \quad (21)$$

$$\bar{E} - e_{nt}^l \geq \frac{15}{60} T_l \beta_n e_{nt}^{rd} + 2T_l \beta_n e_{nt}^{lfd} \quad \forall n \in N, t \in T \quad (22)$$

$$\bar{E} - e_{n,t+1}^l \geq \frac{15}{60} T_l \beta_n e_{nt}^{rd} + 2T_l \beta_n e_{nt}^{lfd} \quad \forall n \in N, t \in T \quad (23)$$



$$\sum_{q=t-UT_g+1}^t s_{gq} \leq u_{gt} \quad \forall g \in G, t \in T \quad (24)$$

$$\sum_{q=t+1}^{t+DT_g} s_{gq} \leq 1 - u_{gt} \quad \forall g \in G, t \in T \quad (25)$$

$$-RR_g \leq q_{g,t+1} - q_{gt} \leq RR_g \quad \forall g \in G, t \in T \quad (26)$$

$$q_{gt} + r_{gt}^u + lf_{ngt}^u \leq \bar{Q}_g u_{gt} \quad \forall g \in G, t \in T \quad (27)$$

$$q_{gt} + r_{gt}^d + lf_{gt}^d \geq \underline{Q}_g u_{gt} \quad \forall g \in G, t \in T \quad (28)$$

$$s_{gt} \geq u_{gt} - u_{g,t-1} \quad \forall g \in G, t \in T \quad (29)$$

$$0 \leq r_{gt}^u \leq \bar{R}_g u_{gt} \quad \forall g \in G, t \in T \quad (30)$$

$$0 \leq r_{ngt}^d \leq \underline{R}_{ng} u_{ngt} \quad \forall g \in G, t \in T \quad (31)$$

$$0 \leq lf_{gt}^u \leq \bar{L}F_g u_{gt} \quad \forall g \in G, t \in T \quad (32)$$

$$0 \leq lf_{gt}^d \leq \underline{L}F_g u_{gt} \quad \forall g \in G, t \in T \quad (33)$$

$$e_{nt}^{cr}, e_{nt}^{dr}, e_{nt}^{ru}, e_{nt}^{rd}, e_{nt}^{lfu}, e_{nt}^{lfd} \geq 0 \quad \forall n \in N, t \in T \quad (34)$$

$$u_{gt} \in \{0, 1\} \quad \forall g \in G, t \in T \quad (35)$$

$$0 \leq s_{gt} \leq 1 \quad \forall g \in G, t \in T \quad (36)$$

$$\sum_{g \in G} (r_{gt}^u) + \sum_{n \in N} (e_{nt}^{ru}) \geq \rho \left( \max_{a \in T: 24-a \geq t} L_{na} + \sum_{n \in N} (\bar{S}_n + \bar{W}_n) \right) \quad \forall t \in T \quad (37)$$

$$\sum_{g \in G} (r_{gt}^d) + \sum_{n \in N} (e_{nt}^{rd}) \geq \rho \left( \max_{a \in T: 24-a \geq t} L_{na} + \sum_{n \in N} (\bar{S}_n + \bar{W}_n) \right) \quad \forall t \in T \quad (38)$$

$$\sum_{g \in G} (lf_{gt}^u) + \sum_{n \in N} (e_{nt}^{lfu}) \geq \eta \sum_{n \in N} L_{nt} + \nu \sum_{n \in N} (S_{nt} + W_{nt}) \quad \forall t \in T \quad (39)$$

$$\sum_{g \in G} (lf_{gt}^d) + \sum_{n \in N} (e_{nt}^{lfd}) \geq \nu \sum_{n \in N} (S_{nt} + W_{nt}) \quad \forall t \in T \quad (40)$$

Variable	Description
$q_{g,t}$	Power generated by generator $g$ in time period $t$
$u_{g,t}$	Binary variable denoting if generator $g$ is operated in period $t$
$s_{g,t}$	Continuous variable denoting if generator $g$ is started in period $t$
$e_{n,t}$	Storage level (energy) at node $n$ at start of period $t$
$e_{nt}^{cr}$	Power charged to storage at node $d$ during period $t$
$e_{nt}^{dr}$	Power discharged from storage at node $d$ during period $t$
$e_{nt}^{ru}$	Power available to provide upreg from storage at node $d$ in period $t$
$e_{nt}^{rd}$	Power available to provide downreg from storage at node $d$ in period $t$
$c_{d,t}^{Charging}$	Binary variable denoting if battery is charging at demand node $d$ in period $t$
$c_{d,t}^{Discharging}$	Binary variable denoting if battery is discharging at demand node $d$ in period $t$
$r_{g,t}^{Up}$	Amount of upreg supplied by generator $g$ in period $t$
$r_{g,t}^{Down}$	Amount of downreg supplied by generator $g$ in period $t$
$\theta_{i,t}$	Voltage angle at node $i$ in time period $t$ (demand or generator)

Set	Description
$N$	Set of all nodes in network
$G$	set of generators
$G_n$	subset of $G$ ; set of generators located at node $n$
$T$	Set of all time periods

## References

- [1] Thomas Ackermann, Goran Andersson, and Lennart Soder. Distributed generation: a definition. *Electric Power Systems Research*, 57(3):195 – 204, 2001.
- [2] US Energy Information Administration. Short Term Energy Outlook. <http://www.eia.gov/forecasts/steo/archives/jan08.pdf>, 2008. [Online; accessed 30-August-2013].
- [3] US Energy Information Administration. Short Term Energy Outlook. <http://www.eia.gov/forecasts/steo/archives/jan13.pdf>, 2013. [Online; accessed 30-August-2013].
- [4] E.H. Allen and M.D. Ilic. Reserve markets for power systems reliability. *Power Systems, IEEE Transactions on*, 15(1):228 –233, Feb. 2000.
- [5] Kenneth J. Arrow, Theodore Harris, and Jacob Marschak. Optimal inventory policy. *Econometrica*, 19(3):250–272, 1951.
- [6] S. Axsäter. *Inventory control*. International series in operations research & management science. Springer, 2006.
- [7] P.P. Barker and R.W. De Mello. Determining the impact of distributed generation on power systems. i. radial distribution systems. In *Power Engineering Society Summer Meeting, 2000. IEEE*, volume 3, pages 1645 –1656, Jul. 2000.
- [8] J.P. Barton and D.G. Infield. Energy storage and its use with intermittent renewable energy. *Energy Conversion, IEEE Transactions on*, 19(2):441 – 448, Jun. 2004.
- [9] G.N. Bathurst and G. Strbac. Value of combining energy storage and wind in short-term energy and balancing markets. *Electric Power Systems Research*, 67(1):1 – 8, 2003.
- [10] R. Bellman, I. Glicksberg, and O. Gross. On the optimal inventory equation. *Management Science*, 2(1):83–104, 1955.
- [11] Richard Bellman. *Dynamic Programming*. Dover Publications, March 2003.
- [12] Dimitri P. Bertsekas. *Dynamic Programming and Optimal Control, Two Volume Set*. Athena Scientific, 2005.
- [13] E. Bitar, R. Rajagopal, P. Khargonekar, and K. Poolla. The role of co-located storage for wind power producers in conventional electricity markets. In *American Control Conference (ACC), 2011*, pages 3886 –3891, Jun. 2011.
- [14] Severin Borenstein, James Bushnell, Christopher R. Knittel, and Catherine Wolfram. Inefficiencies and market power in financial arbitrage: A study of california’s electricity markets. *The Journal of Industrial Economics*, 56(2):347–378, 2008.
- [15] John Bjørnar Bremnes. Probabilistic wind power forecasts using local quantile regression. *Wind Energy*, 7(1):47–54, 2004.

- [16] C. Bueno and J.A. Carta. Wind powered pumped hydro storage systems, a means of increasing the penetration of renewable energy in the canary islands. *Renewable and Sustainable Energy Reviews*, 10(4):312 – 340, 2006.
- [17] J. Bushnell, S. M. Harvey, and B. F. Hobbs. Opinion on pay-for-performance regulation, Mar. 2012.
- [18] California ISO. Regulation energy management draft final proposal, Jan. 2011.
- [19] Thomas Carr'. Deriving wecc carbon co2 emissions and capital costs. Technical Advisory Subcommittee Meeting, Phoenix, AZ, 2008.
- [20] J.M. Carrasco, L.G. Franquelo, J.T. Bialasiewicz, E. Galvan, R.C.P. Guisado, Ma.A.M. Prats, J.I. Leon, and N. Moreno-Alfonso. Power-electronic systems for the grid integration of renewable energy sources: A survey. *Industrial Electronics, IEEE Transactions on*, 53(4):1002 –1016, Jun. 2006.
- [21] Mike Chen, In-Koo Cho, and Sean P. Meyn. Reliability by design in distributed power transmission networks. *Automatica*, 42(8):1267 – 1281, 2006.
- [22] Western Electricity Coordinating Council. Transmission Expansion Planning Program Committee 2009 Study Program Results Report. [http://www.wecc.biz/committees/BOD/TEPPC/External2009\\_TEPPC\\_StudyResults\\_Report.pdf](http://www.wecc.biz/committees/BOD/TEPPC/External2009_TEPPC_StudyResults_Report.pdf), 2010. [Online; accessed 19-July-2013].
- [23] Gerard Debreu. A social equilibrium existence theorem. *Proceedings of the National Academy of Sciences of the United States of America*, 38(10):886–893, Oct. 1952.
- [24] Paul Denholm and Ramteen Sioshansi. The value of compressed air energy storage with wind in transmission-constrained electric power systems. *Energy Policy*, 37(8):3149 – 3158, 2009.
- [25] DOE. Energy storage program planning document, Feb. 2011.
- [26] Easan Drury, Paul Denholm, and Ramteen Sioshansi. The value of compressed air energy storage in energy and reserve markets. *Energy*, 36(8):4959 – 4973, 2011. <ce:title>PRES 2010</ce:title>.
- [27] J Eyer and G Corey. Energy Storage for the Electricity Grid: Benefits and Market Potential Assessment Guide A Study for the DOE Energy Storage Systems Program. Sandia Report SAND2010-0815, 2010. [Online; accessed 27-March-2012].
- [28] Natalia Fabra, Nils-Henrik von der Fehr, and David Harbord. Modeling electricity auctions. *The Electricity Journal*, 15(7):72 – 81, 2002.
- [29] Ky Fan. Fixed-point and minimax theorems in locally convex topological linear spaces. *Proceedings of the National Academy of Sciences of the United States of America*, 38(2):121–126, 1952.
- [30] Deqiang Gan and E. Litvinov. Energy and reserve market designs with explicit consideration to lost opportunity costs. *Power Systems, IEEE Transactions on*, 18(1):53 – 59, Feb. 2003.
- [31] T.W. Gedra. On transmission congestion and pricing. *Power Systems, IEEE Transactions on*, 14(1):241 –248, feb 1999.
- [32] I. L. Glicksberg. A further generalization of the Kakutani fixed point theorem, with application to nash equilibrium points. *Proceedings of the American Mathematical Society*, 3(1):170–174, 1952.

- [33] John C. Harsanyi. Games with incomplete information played by "Bayesian" players, i-iii. part i. the basic model. *Management Science*, 14(3):159–182, 1967.
- [34] F.B. Hildebrand. *Introduction to Numerical Analysis*. Dover Books on Advanced Mathematics. Dover Publications, 1987.
- [35] H. Ibrahim, A. Ilinca, and J. Perron. Energy storage systems - characteristics and comparisons. *Renewable and Sustainable Energy Reviews*, 12(5):1221 – 1250, 2008.
- [36] Electric Power Research Institute. Electricity Energy Storage Technology Options A White Paper Primer on Applications, Costs and Benefits. <http://www.epri.com/abstracts/Pages/ProductAbstract.aspx?ProductId=000000000001020676>, 2010. [Online; accessed 30-August-2013].
- [37] Jae Ho Kim and Warren B. Powell. Optimal energy commitments with storage and intermittent supply. *Operations Research*, 59(6):1347–1360, November/December 2011.
- [38] Magnus Korpaas, Arne T. Holen, and Ragne Hildrum. Operation and sizing of energy storage for wind power plants in a market system. *International Journal of Electrical Power & Energy Systems*, 25(8):599 – 606, 2003.
- [39] R.H. Lasseter. Microgrids. In *Power Engineering Society Winter Meeting, 2002*, volume 1, pages 305 – 308, 2002.
- [40] S. Lemofouet and A. Rufer. A hybrid energy storage system based on compressed air and supercapacitors with maximum efficiency point tracking (mept). *Industrial Electronics, IEEE Transactions on*, 53(4):1105 –1115, june 2006.
- [41] Nils Löhndorf and Stefan Minner. Optimal day-ahead trading and storage of renewable energies: an approximate dynamic programming approach. *Energy Systems*, 1:61–77, 2010. 10.1007/s12667-009-0007-4.
- [42] V Loose. Quantifying the Value of Hydropower in the Electric Grid: Role of Hydropower in Existing Markets. Sandia Report SAND2011-1009, 2011.
- [43] John F. Nash. Non-cooperative games. *The annals of mathematics*, 54(2):286–295, Sep. 1951.
- [44] A. Papavasiliou. *Coupling Renewable Energy Supply with Deferrable Demand*. PhD thesis, University of California, Berkeley, 2011.
- [45] Mukund R. Patel. *Wind and solar power systems: design, analysis, and operation*. CRC Press, 2006.
- [46] Nicholas C. Petruzzi and Maqbool Dada. Pricing and the newsvendor problem: A review with extensions. *Operations Research*, 47(2):183–194, 1999.
- [47] PIER. 2020 strategic analysis of energy storage in California, Dec. 2010.
- [48] P. Pinson, C. Chevallier, and G.N. Kariniotakis. Trading wind generation from short-term probabilistic forecasts of wind power. *Power Systems, IEEE Transactions on*, 22(3):1148 –1156, Aug. 2007.
- [49] E.L. Porteus. *Foundations of stochastic inventory theory*. Stanford Business Books. Stanford Business Books, an imprint of Stanford University Press, 2002.

- [50] J.E. Price and J. Goodin. Reduced network modeling of wecc as a market design prototype. *Power and Energy Society General Meeting, IEEE*, 24(2):1039–1050, 2011.
- [51] P.F. Ribeiro, B.K. Johnson, M.L. Crow, A. Arsoy, and Y. Liu. Energy storage systems for advanced power applications. *Proceedings of the IEEE*, 89(12):1744 –1756, Dec. 2001.
- [52] J. B. Rosen. Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica*, 33(3):520–534, 1965.
- [53] Tim Roughgarden. *Selfish Routing and the Price of Anarchy*. The MIT Press, 2005.
- [54] Fred C. Schweppe, Michael C. Caramanis, Richard D. Tabors, and Roger E. Bohn. *Spot pricing of electricity*. Kluwer Academic Publishers, Boston, MA, 1988.
- [55] Edward A. Silver, David F. Pyke, and Rein Peterson. *Inventory Management and Production planning and Scheduling*. Wiley, 3rd edition, 1998.
- [56] R. Sioshansi and P. Denholm. The value of plug-in hybrid electric vehicles as grid resources. *The Energy Journal*, 31(3):1–23, 2010.
- [57] Ramteen Sioshansi, Paul Denholm, Thomas Jenkin, and Jurgen Weiss. Estimating the value of electricity storage in pjm: Arbitrage and some welfare effects. *Energy Economics*, 31(2):269 – 277, 2009.
- [58] A. Skartveit and J.A. Olseth. The probability density and autocorrelation of short-term global and beam irradiance. *Solar Energy*, 49(6):477 – 487, 1992.
- [59] J. Usaola, O. Ravelo, G. Gonzalez, F. Soto, C. Davila, and B. Diaz-Guerra. Benefits for wind energy in electricity markets from using short term wind power prediction tools—A simulation study. *Wind Eng.*, 28(1):119–128, Feb. 2004.
- [60] S.R. Vosen and J.O. Keller. Hybrid energy storage systems for stand-alone electric power systems: optimization of system performance and cost through control strategies. *International Journal of Hydrogen Energy*, 24(12):1139 – 1156, 1999.
- [61] H. L. Willis. *Distributed Power Generation: Planning and Evaluation*. CRC Press, 2000.
- [62] T Wu, M. Rothleder, Z. Alaywan, and A.D. Papalexopoulos. Pricing energy and ancillary services in integrated market systems by an optimal power flow. *IEEE Transactions on Power Systems*, 19(1):339–347, 2004.
- [63] Tongxin Zheng and E. Litvinov. Ex post pricing in the co-optimized energy and reserve market. *Power Systems, IEEE Transactions on*, 21(4):1528 –1538, Nov. 2006.
- [64] Yangfang Zhou, Alan Andrew Scheller-Wolf, Nicola Secomandi, and Stephen Smith. Managing wind-based electricity generation with storage and transmission capacity. 2012. Available at SSRN: <http://dx.doi.org/10.2139/ssrn.1962414>.

## 4.6 Advanced Energy Storage – Financing Mechanisms

### Introduction

The value of PV and storage spans multiple services and beneficiaries. For the site owner, storage provides backup power and savings on retail rates through reduced demand charges and energy charges. For a utility, storage can reduce congestion and defer transmission and distribution equipment upgrades. For a Load Balancing Authority such as CAISO, storage can offer additional capacity for resource adequacy planning, wholesale energy arbitrage, and ancillary services like voltage regulation that all improve grid reliability and reduce carbon emissions. Despite the clearly identified value, there exist few viable technologies and market mechanisms that enable the beneficiaries to capture the value of combined PV and storage.

SolarCity has demonstrated that a zero-down, cash-flow-positive finance mechanism can enable rapid adoption and deployment of PV. SolarCity's finance products direct private sector tax equity investments toward financing PV system installations. When structured appropriately, these investment mechanisms enable many host customers to benefit from PV for no upfront cost, with an accompanying monthly finance payment that may be lower than their offset utility bill. SolarCity found that one of the key barriers to deploying PV was a high upfront cost. These financing mechanisms eliminate or reduce this barrier. Since 2008 when SolarCity launched its first residential finance product, the proportion of financed residential systems in California has grown from 0% to more than 65%, according to the GTM Research Solar Market Insight Report from Q4 2012. Overall, the US residential market has grown substantially installing more in one quarter in 2012 than was installed in all of 2009.<sup>15</sup> In a related development, the prevalence of solar financing models has coincided with a dramatic increase in adoption of solar in lower and middle income areas. A June 2012 study from the California Public Utilities Commission reported a 364% increase in applications since the program's inception from households in zip codes with median incomes of less than \$50,000, and a 445% increase in applications from households in zip codes with median incomes between \$50,000 and \$75,000. The report notes that *"the upward trend in CSI participation in lower and middle income areas is likely due to a sharp increase in third party owned systems that have received CSI incentives. Third party ownership models, such as solar leases and power purchase agreements (PPAs), allow households who cannot afford to own a PV system to go solar."*<sup>16</sup>

Many of the lessons from PV financing could be applied to create a successful finance program for distributed storage installations. Similar structures, contracts, and sales and marketing techniques could apply to storage. As a result, the right financing models could accelerate the deployment of storage systems just as they have accelerated the adoption of PV.

---

<sup>15</sup> <http://www.greentechmedia.com/articles/read/U.S.-Solar-Market-Spikes-with-742-MW-in-Solar-Installations-in-Q2-2012/>

<sup>16</sup> <http://www.cpuc.ca.gov/NR/rdonlyres/0C43123F-5924-4DBE-9AD2-8F07710E3850/0/CASolarInitiativeCSIAnnualProgAssessmtJune2012FINAL.pdf>

## Background

Distributed energy storage has enormous potential to shape the future of high efficiency energy management. The Electric Power Research Institute (EPRI) published a study in December 2010<sup>17</sup> that identified many of the potential benefits of energy storage:

Value Chain	Benefit
End User	Power Quality
	Power Reliability
	Retail Time of Use Energy Charges
	Retail Demand Charges
Distribution	Voltage Support
	Defer Distribution Investment
	Distribution Losses
Transmission	VAR Support
	Transmission Congestion
	Transmission Access Charges
	Defer Transmission Investment
System	Local Capacity
	System Capacity
	Renewable Energy Integration
Independent System Operator (ISO) Markets	Fast Regulation (1 hr)
	Regulation (1 hr)
	Spinning Reserves
	Non-Spinning Reserves
	Black Start
	Price Arbitrage

Despite these benefits, the high upfront cost of new technologies often prohibits wide-scale deployment. As a result, the ability to inexpensively finance new technologies is critical in the path to broader adoption.

Grid interactive storage is a less developed industry than PV and there is significantly less data on the performance of storage systems over time. Financing new technologies requires a strong understanding of future cash flows and the risks surrounding the collectability of those cash flows. As a result, the first step toward financing a technology should involve gaining a firm understanding of the technology in order to quantify the value and identify of risks. The following discussion attempts to quantify the value proposition and identify the risks for firm PV.

## Discussion

The value of distributed energy storage can be grouped into three main categories, as summarized in the following table:

---

<sup>17</sup> Electric Power Research Institute. *Electricity Energy Storage Technology Options: A White Paper Primer on Applications, Costs, and Benefits*. Technical Update, December 2010.



Value Proposition	Benefit
Renewable Integration	Decreased reliance on other energy sources, through a sustainable and free power source (the sun)
Distributed Infrastructure Support	Grid transmission and distribution support from centrally managed systems able to charge, store, and provide power to the grid
Local Energy Management	Reliable backup power for an energy user with the potential to shift time of use energy and demand charges

In order to obtain financing for energy storage, we need to first understand the timing and amount of benefits generated. In addition, financier concerns stem from the certainty of repayment, which requires a risk-adjusted reward. As a result, we need to assess the certainty that the potential benefits will be realized at the amount and time anticipated.

### Renewable Integration

The monetary value of renewable integration is set by the market demand for clean energy. In other words, the monetary value of renewable integration is set by a simple question: “What would an individual, a utility, a corporation, or any other entity pay for clean energy?” Environmental motivations, while important to the demand for clean energy sources, have fallen short in driving renewable integration in areas where the cost of adoption is economically challenging (i.e. because other energy sources are currently less expensive and the cost of pollution and/or the depletion of natural resources are incorrectly priced, or the myriad benefits of distributed renewable generation are not properly accounted for). In response, many states have created clean power mandates, also referred to as Renewable Portfolio Standards (RPS), in an attempt to increase the demand for renewable integration.

To facilitate a pricing and trading market for renewable energy, states with RPS mandates often award Renewable Energy Certificates (RECs) based on the generation of power from an approved energy source and provide liquidity for an otherwise illiquid market. While helpful in spurring renewable integration, RPS programs exist only in select states and REC bidding prices have been historically volatile, which impairs the ability to finance a new technology because the future cash flow is difficult to predict.

**Distributed Infrastructure Support** – The EPRI study on energy storage referenced above attempted to quantify the potential value of infrastructure support. Financing distributed infrastructure support has proven particularly challenging because the regulatory environment has not yet evolved to allow, utilize, or efficiently price the potential infrastructure benefits for storage.

**Local Energy Management** – Local energy management is arguably easier to finance than the benefits of renewable integration or distributed infrastructure support, since a storage provider could enter into a contract with a building owner to provide energy management services over a specified time. The pricing for a local energy management contract would be based on determining the potential value to the energy user from two main benefits: 1) energy management to reduce time-of-use (TOU) energy

costs and demand costs and 2) the ability for energy storage devices to provide power when the grid is otherwise down. The resulting contract with the building owner could, in turn, be financed. In this scenario, the credit-worthiness of the building owner could be an indicator of the certainty of collectability from a financing perspective. Additionally, the storage installation could serve as collateral in the event of default.

Regardless of the value proposition, additional uncertainties may impair the ability to finance storage. These uncertainties are inherent in emerging categories where early adopters and financiers are often required to predict which technologies will ultimately succeed or fail to become both commercially viable and widely adopted. Additionally, newer technologies have limited performance history, which inhibits the developer's ability to prove that the technology will function as intended.

In summary, the challenges of financing storage are as follows:

Value Proposition	Financing Challenge
Renewable Integration	<ul style="list-style-type: none"> <li>• Amount and collectability of REC income is difficult to predict due to volatility of bidding prices</li> <li>• Many states do not have regulations that require renewable integration</li> <li>• The technology may not perform as intended</li> <li>• The industry may adopt a different technology</li> </ul>
Distributed Infrastructure Support	<ul style="list-style-type: none"> <li>• Regulatory environment needs to change before the infrastructure benefits can be efficiently realized and priced/financed</li> <li>• The technology may not perform as intended</li> <li>• The industry may adopt a different technology</li> </ul>
Local Energy Management	<ul style="list-style-type: none"> <li>• Determining the value of local energy management is difficult due to the lack of a large, well-developed and transparent marketplace</li> <li>• The technology may not perform as intended</li> <li>• The industry may adopt a different technology</li> </ul>

An effective finance structure can overcome these barriers by allocating risk and return and distributing incentives and benefits to the optimal parties.

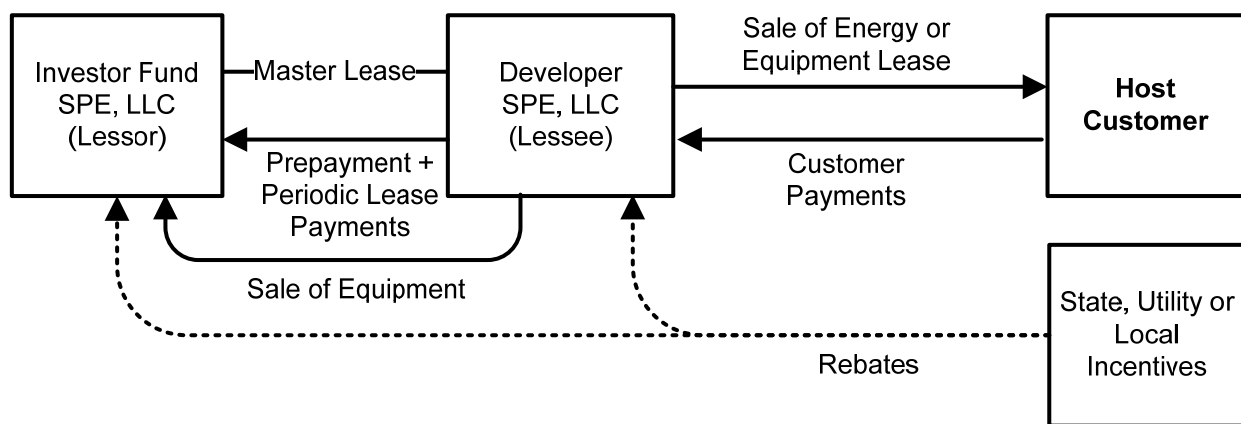
As with PV, there are a variety of potential incentives (in addition to the value identified above) that could enhance the financial incentive to deploy storage systems. These include state and local utility cash rebate programs like the Self-Generation Incentive Program in California, which pays eligible storage technologies up to \$1.80/Watt installed in 2013.<sup>18</sup> In addition, the owners of these systems may capture depreciation benefits and accelerated depreciation for some classes of assets.

There are a number of potential methods of financing high-cost assets like batteries that otherwise may be difficult to finance on a corporate balance sheet, or may be more attractive to finance independently.

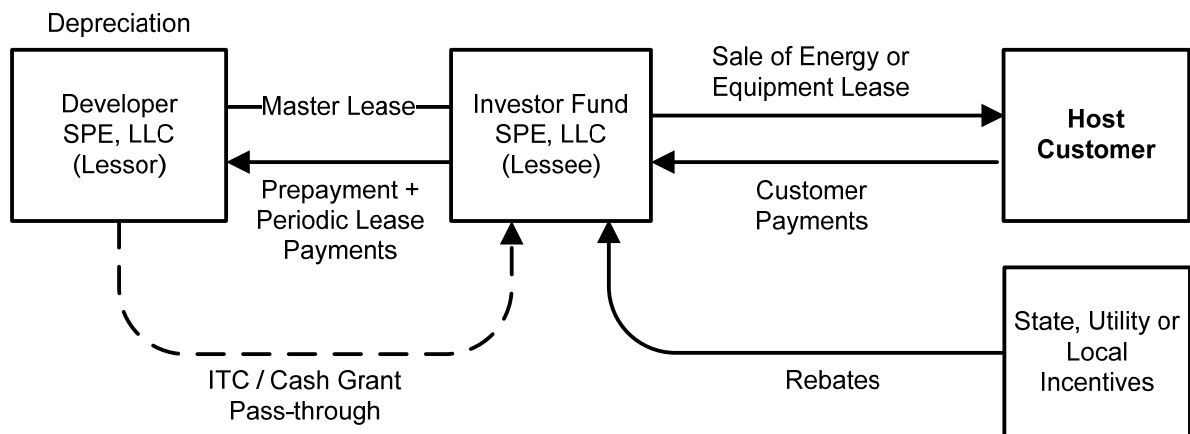
<sup>18</sup> [http://www.pge.com/includes/docs/pdfs/shared/newgenerator/selfgeneration/SGIP\\_Handbook\\_2012.pdf](http://www.pge.com/includes/docs/pdfs/shared/newgenerator/selfgeneration/SGIP_Handbook_2012.pdf)

Typically, some combination of project equity and debt is used to fund installations of assets like batteries. Where tax incentives exist, tax equity may be tapped as well. The three primary project equity finance structures used to finance solar photovoltaic projects in the US today are the sale leaseback, the inverted lease, and the partnership flip.

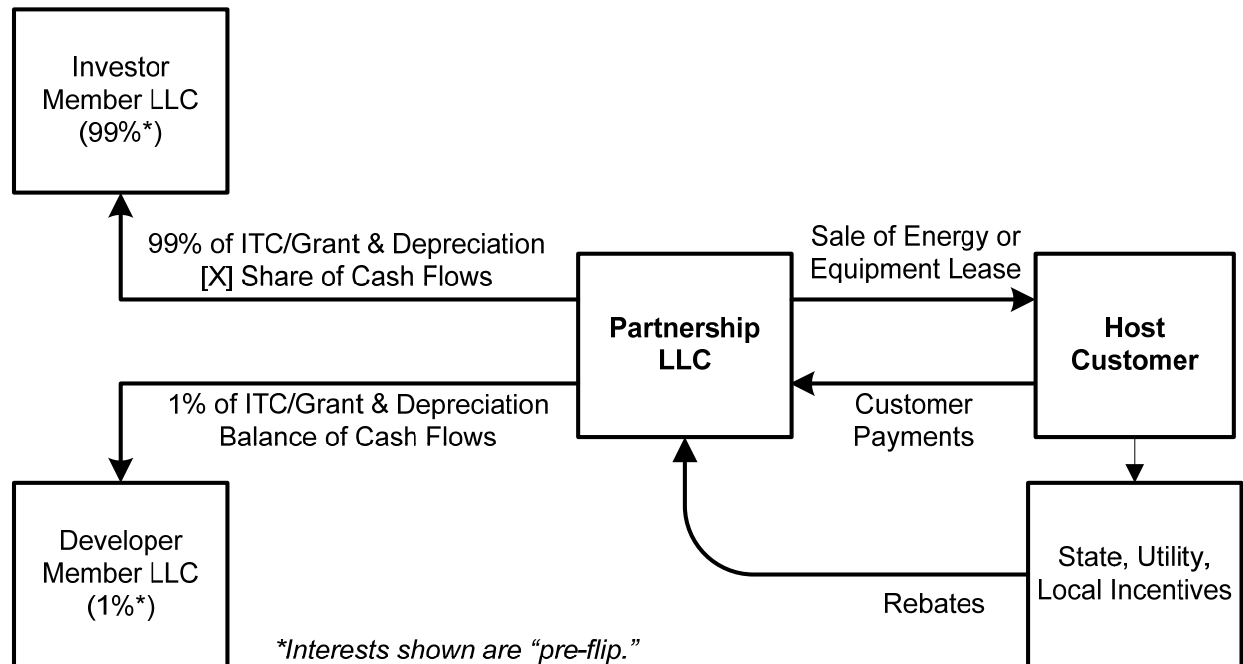
**Sale Leaseback** - In a sale leaseback structure, the project developer installs the system and then sells it back to the equity investor. The equity investor then leases the system back to the developer who uses some of the customer payments to pay rent to the investor, as lessor. The investor is able to claim any tax or depreciation benefits or rebates that accrue. At the end of the lease term, the investor can either retain ownership of the system or, if the Developer, as lessee, elects to exercise its purchase option, sell it to the developer at the market value.



**Inverted Lease** - In an inverted lease, the developer installs and owns the system, which it leases to the investor, who in turn sub-leases the system to customers. The developer, as lessor, retains any depreciation benefits as the owner of the asset. The investor makes master lease payments to the developer and keeps the rent payments that come from the end customer. If there is an applicable tax incentive, it can be passed through to the investor.



**Partnership Flip** - In a partnership flip, the investor and developer create a partnership through which they co-own the asset. The partners agree on the proportion of cash and tax benefits that accrue to each party and a target rate of return required by the investor. Often the investor retains a 99% interest in the revenue and tax credits as well a share of cash until achieving a pre-negotiated target return. Once the investor reaches the target return, the partnership interest “flips” so that the developer gets typically 95% of the benefits thereafter with an option to purchase the investor’s remaining 5% interest at the market value. If the purchase option is exercised, the developer may retain full ownership of the asset.



The selection of the best structure for a given project depends on many factors. These can include the credit rating of the end customer, the technology’s track record in existing deployments; the efficacy of utilizing the technology in delivering proven economic value to customers; market conditions; availability of specific sources of financing; and the applicable technology’s eligibility for tax credits and/or rebates. All of the above equity structures may be supplemented with or replaced by some form of debt. Clearly there are many viable options to choose from.

## Conclusion

The combination of PV and grid interactive storage can achieve substantial cost savings for utilities and end customers—and reduce carbon emissions to a far greater degree than either PV or storage could achieve on their own—while helping ease the strains on aging utility infrastructure. The key to unlocking these benefits is overcoming the barriers to adoption including upfront costs. Financing can enable broader adoption rates of a technology like storage with high upfront capital costs by allowing customers to align periodic payments over the creation of benefit. The same innovative finance

mechanisms that have enabled recent growth in the distributed solar PV industry may well ease growth in deployments of distributed energy storage systems.

## Conclusions

The State of California is undergoing a transformation of its energy generation infrastructure. The CSI program and other policy initiatives have created a strong market for technologies that reduce the carbon dependence of the energy economy and a particularly vibrant solar PV market. However, as grid penetrations of PV increase it is important to understand ways to “firm” and increase the quality of the energy that PV delivers to the grid. Advanced energy storage paired with new or existing PV generation is likely to be one of the primary tools to this end.

The objective of this grant was to develop advanced battery storage products, demonstrate the integration of these products with existing solar PV assets, analyze the value streams that these dual systems could provide, and identify market mechanisms by which this value can be accessed. These goals were laid out as follows and divided into four tasks with various subtasks.

1. (Task 4) Establish extent of PV variability to determine energy storage system needs
2. (Task 2) Demonstrate a cost effective pairing of PV and energy storage for load shifting, demand reduction, and conventional ancillary services
3. (Task 3) Show that advanced, distributed, PV-coupled, grid-interactive storage solutions will reduce cost and carbon emissions and improve grid reliability and security
4. (Task 4) Identify utility retail and ISO wholesale rate structures, tariffs, and market mechanisms that will be necessary to bring combined PV and storage to new markets
5. (Task 3) Reduce the cost of installation (balance of system and labor)
6. (Task 3) Streamline the interconnection process for “firm” PV systems

The research in Task 4 suggests that PV variability provides net benefit to the grid at moderate penetrations and that even at high penetrations, the load/generations-shifting capabilities of the advanced energy storage systems increase the avoided energy cost and avoided capacity cost that are already benefits of added PV generation. The models in Task 4.4 also show that the system operating cost is reduced as the penetration of energy storage systems participating in the regulation markets is increased (decreased generator starts). It was also noted that under some conditions, increased penetration of energy storage could increase the carbon intensity of the system as a whole.

The development and deployment Tasks were also successful. SolarCity and Tesla were able to design, develop, and install both residential and commercial advanced lithium ion products. Throughout the process there were many insights gathered on important product specifications, code requirements, installation process, and customer feedback.

All tasks were all completed and delivered benefits to the parties involved. SolarCity and Tesla have since built more than one hundred residential systems and more than 10 commercial systems. Overall, the grant was quite successful and has enabled informed guidance in the various policy and regulatory settings that are currently determining the future of paired PV and energy storage products. The continuing evolution of tariffs, technical requirements, market structure, and regulatory environments that guide the operation of the energy resources of California will be critical to the success of energy storage, and ultimately in the state meeting its long term energy goals.

## Next Steps

A great next step for the technology would be a larger pilot focusing more specifically on the advanced use cases of behind the meter energy storage, both on a site-specific and aggregated basis. This would include deploying and dispatching customer-sited storage systems to provide retail services, like peak demand shaving and demand response, but also demonstrating the ability for grid operators to use these systems to provide grid facing services, including utilization to address capacity needs, frequency regulation, voltage regulation, and ramping services.

Storage is a uniquely versatile resource. However, fully exploiting this potential will require a certain degree of comfort with the capabilities of the technology and better alignment with the regulatory framework within which these technologies are deployed. Advanced use case pilots can play an invaluable role in facilitating this.