Real-Time, Multi-Service Operation of Grid-Scale Energy Storage using Model Predictive Control

Kevin Moy

kmoy14@stanford.edu

June 5, 2021

Abstract

Energy storage systems (ESSs) are a critical part of the renewable-fueled, sustainable energy grid of the future. ESSs can increase their value by dispatching (charging and discharging) to the grid to provide multiple grid services. This paper presents a framework to non-simultaneously provide two separate grid services of energy arbitrage and peak shaving in real-time using model predictive control (MPC), demonstrating the trade-off between MPC horizon and computational power needed for real-time control.

Video link: `https://youtu.be/Q9qfMK28X0g`

1 Introduction

Energy storage systems (ESSs) are a critical part of the renewable-fueled, sustainable energy grid of the future. Currently, ESSs are used in such renewable grid systems by dispatching (charging and discharging) to the grid to provide services which support a variety of different stakeholders, including utilities, transmission operators, and utility customers [4]. Generally, these services can be divided into "behind-the-meter" (BTM) services, which support customers, e.g. providing power during a blackout, and "front-of-meter" (FOM) services, which support the grid at large, e.g. regulating grid voltage. Currently, a single energy storage asset can only participate (provide and be compensated for) BTM or FOM services, but not both.

However, ESSs are still expensive, and many applications do not require energy storage dispatch at all times. From this, many operators of ESSs seek to use the ESSs in "value-stacking"; that is, using a singular energy storage resource for multiple different services/grid applications. Value-stacking is supported by federal energy policy, via FERC Order 841, which directed transmission grid operators to provide means for energy storage to participate in both BTM and FOM services, [1] and FERC Order 2222, which broadens and strengthens the ruling in the previous Order. [2] It is also supported by California legislation, via the Multi-Use Application framework [3]. This paper aims

[1]Federal Energy Regulatory Commission, Order 841, Effective Date 06/04/2018. Available: https://www.federalregister.gov/d/2018-03708

[2]Federal Energy Regulatory Commission, Order 2222, Effective Date 10/29/2020. Available: https://www.federalregister.gov/d/2020-20973

[3]https://www.utilitydive.com/news/california-regulators-first-to-allow-multiple-revenue-streams-for-energy-st/516927/

to demonstrate the real-time control of an ESS performing value-stacking using model-predictive control (MPC) for the grid services of peak shaving and energy arbitrage.

2 Related Work

The literature presents a wealth of studies for optimal control of energy storage. Canonically, the effort to determine the optimal dispatch of energy storage without real-time control (that is, with full access to all necessary present, past, and future data is a mixed-integer linear programming (MILP) problem, and in the simplest case, can be reduced to a linear program. In [6], the single objective of bill minimization was used and formulated as an MILP, and a daily optimization vs. weekly optimization was considered for a distributed energy system, which included energy storage. Other studies focus on multi-use frameworks for energy storage. In [8], a single energy storage was proven to exhibit "superlinear" gains by value-stacking both peak-shaving and frequency regulation. However, this paper incorrectly applies the simultaneous benefits of both grid services at once, in direct violation of FERC Order 841. In terms of real-time optimal control, model predictive control (MPC) is the predominant method. MPC is often used for multi-stage optimization; in one paper, a single system is used for different grid services that operate on two different time scales [3], while another paper uses MPC as the inner loop optimization with an outer loop sizing optimization to simultaneously determine the optimal control and sizing for the system [1]. In [2], the MPC is coupled with a physics-based model for lithium-ion batteries, demonstrating the effects of including a penalty function based on the model-derived degradation into the optimization cost function.

3 Problem Statement

This project focuses on the multiple services that energy storage can provide, here simplified into the case of two services. First, the BTM application of peak-shaving, or dispatching the energy storage to reduce the utility bill, and second, the FOM application of energy arbitrage, or "buying low/selling high" on the energy marketplace. The control problem is then to decide whether to dispatch in peak shaving, directly supporting local loads, or dispatch into energy arbitrage, participating in the energy market, or neither, for the goal of maximizing value (peak-shaving savings + energy arbitrage revenue) over a given time period of one week.

4 Approach

4.1 System Definition

The system used in this analysis which allows for such behavior is shown in Figure 1, with power flows labelled using Table 1. For peak shaving, the battery dispatches against the time-of-use (TOU) electricity pricing, discharging to meet the load demand (as measured by the utility meter) when the TOU price is high. For energy arbitrage, the battery dispatches against the locational marginal price (LMP) of energy, charging/discharging directly from/to the grid when the price of energy is low/high, respectively. The system cannot perform both peak shaving and energy arbitrage (the blue or red arrows) simultaneously, nor charge or discharge simultaneously. Furthermore, the system must prevent backfeeding, or the flow of energy back through the energy to the electric grid,

Table 1: List of defined variables in the problem description.

Variable	Name	Units
h	Period of the data in fractions of an hour	unitless
T_w	Number of 15-minute periods in one week	unitless
H	Horizon of MPC	hours
T_H	Number of 15-minute periods in horizon	unitless
$l(t)$	Load, total, at time t	kW
$TOU(t)$	TOU tariff energy cost	\$/kWh
$LMP(t)$	LMP energy price	\$/kWh
P_{nom}	Maximum power output of the ESS	kW
E_{nom}	Maximum energy stored in the ESS	kWh
$E(t)$	Energy stored in the ESS at time t	kWh
$P_{g,ess}(t)$	Power flowing from grid to ESS at time t	kW
$P_{ess,l}(t)$	Power flowing from ESS to load at time t	kW
$P_{g,l}(t)$	Power flowing from grid to load at time t	kW
$g(t)$	Total power flowing from grid at time t for TOU	kW
$d_{TOU}(t)$	Power discharged by the ESS at time t for TOU	kW
$c_{TOU}(t)$	Power charged to the ESS at time t for TOU	kW
$d_{LMP}(t)$	Power discharged by the ESS at time t for LMP	kW
$c_{LMP}(t)$	Power charged to the ESS at time t for LMP	kW
$b_{c,TOU}(t)$	Binary decision variable for TOU charging	unitless
$b_{d,TOU}(t)$	Binary decision variable for TOU discharging	unitless
$b_{c,LMP}(t)$	Binary decision variable for LMP charging	unitless
$b_{d,LMP}(t)$	Binary decision variable for LMP discharging	unitless

which would result in a negative utility bill. Such behavior is not currently allowed under FERC Order 841.

The system is sized equivalent to a Tesla Powerwall, rated at $P_{nom} = 5$kW continuous power and $E_{nom} = 14$kWh rated energy[4], and assume that the energy storage has a dispatching efficiency of 100%, and that the continuous power rating of 5 kW holds as the power limit for both charging and discharging.

4.2 Data Description

Peak shaving requires time-series data of the building load and the TOU pricing. The load data are obtained from the Pecan Street Dataport database[5] for a house in San Diego. The relevant TOU plan (TOU-DR1) is also selected from the San Diego Gas & Electric rate tariff[6]. All data are obtained at 15-minute intervals (i.e. $h = 0.25$) for the dates of July 8th 00:00 to June 30th 23:45 of the next year, representing one year of data. Energy arbitrage requires the LMP for every hour in the day. The LMP data are collected from the COVID-EMDA datahub, using CAISO data for a node in the same location as the TOU and load data [7]. The LMP data span the period beginning

[4]https://www.tesla.com/powerwall
[5]https://www.pecanstreet.org/dataport/
[6]https://www.sdge.com/whenmatters

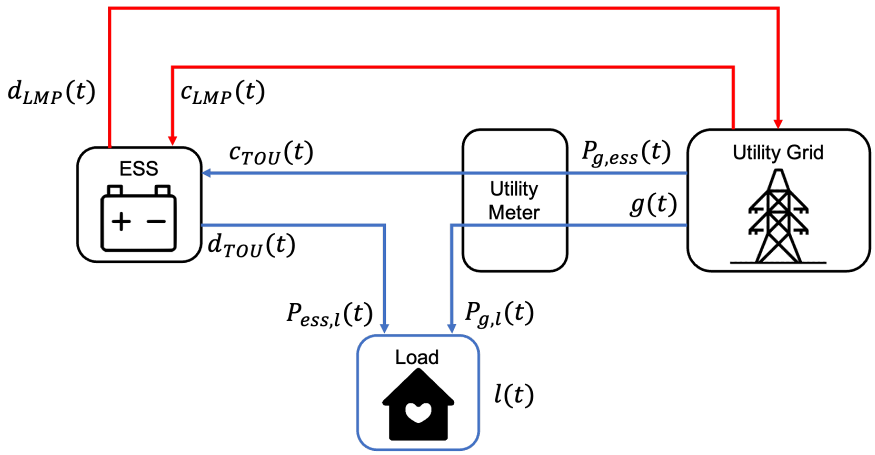


Figure 1: Power flow diagram of system in this paper. Arrows denote direction of power flow. The ESS can charge and discharge via TOU to offset the energy consumed by the load from the grid, as measured via the utility meter. Likewise, it can also charge and discharge directly to the grid, buying and selling energy via LMP energy arbitrage, but cannot perform both simultaneously.

July 8th 2018, 00:00 and ending 30 June 2019, 23:45, for the URBAN-N005 node in San Diego. This data is then upsampled to 15-minute intervals to match those of the tariff and load data.

4.3 Control Methodology

This paper follows the existing state-of-the-art in the literature of using model predictive control as the basis to derive the optimal control of the energy storage system. The cost function is defined as follows:

$$J = \sum_{i=0}^{T_H} h[LMP(i)(d_{LMP}(i) - c_{LMP}(i)) - TOU(t)g(t)] \tag{1}$$

Where the first term corresponds to the revenue from LMP market dispatch, and the second to the (negative) electric bill from TOU. This cost function is maximized for each MPC time step i over a fixed horizon H for a total of $h * H = T_H$ time steps.

The constraints for the MPC correspond to the physical constraints of the system, for known $E_{nom}, P_{nom}, TOU, LMP$, and l. First, the stored energy in the battery cannot exceed its limits at all time steps i

$$0 \le E(i) \le E_{nom} \tag{2}$$

Next, at all time steps i, the power flow between the grid, load, and ESS must be respected

$$c_{TOU}(i) = P_{g,ess}(i) \tag{3}$$

$$d_{TOU}(i) = P_{ess,l}(i) \tag{4}$$

$$g(i) = P_{g,ess}(i) + P_{g,l}(i) \tag{5}$$

$$l(i) = P_{g,l}(i) + P_{ess,l}(i) \tag{6}$$

In addition, a constraint will also be added to ensure that the energy storage system cannot simultaneously dispatch for both peak shaving and energy arbitrage, and that the dispatch power remains within the ESS rated power. This is accomplished via the use of binary variables representing when a particular dispatch is active, e.g. $b_{c,TOU}(i) = 1$ when the ESS charges for peak shaving and 0 otherwise at time step i. The use of binary variables to prevent simultaneous discharge and charge of an ESS is well-established in the literature [5], and is extended here to prevent simultaneous dispatch across multiple applications. For all i,

$$0 \leq d_{TOU}(i) \leq b_{d,TOU}(i) P_{nom} \tag{7}$$

$$0 \leq c_{TOU}(i) \leq b_{c,TOU}(i) P_{nom} \tag{8}$$

$$0 \leq d_{LMP}(i) \leq b_{d,LMP}(i) P_{nom} \tag{9}$$

$$0 \leq c_{LMP}(i) \leq b_{c,LMP}(i) P_{nom} \tag{10}$$

$$b_{d,TOU}(i) + b_{c,TOU}(i) + b_{d,LMP}(i) + b_{c,LMP}(i) = 1 \tag{11}$$

The evolution of energy stored with dispatch must also be respected for time steps $i > 0$:

$$E(i) = h(c_{LMP}(i-1) + c_{TOU}(i-1)) - h(c_{LMP}(i-1) + c_{TOU}(i-1)) + E(i-1) \tag{12}$$

Finally, the stored energy must be initialized to some value, and is constrained to end at that same value at the end of the horizon to ensure that the ESS does not run out of energy:

$$E(0) = E_0 \tag{13}$$

$$E(T_H) = E_0 \tag{14}$$

The value of E_0 is one of the inputs to the MPC algorithm, and serves to preserve continuity of the ESS state between time steps.

4.4 MPC Algorithm

The algorithm for one iteration of MPC can be found in Algorithm 1. This algorithm is carried out at each time step $t = 0, 1, 2, ..., T_W$. At each successive time t, each iteration of MPC of horizon H uses the stored energy from the previous time step (i.e. setting `E_0 = E_1`) and the load, TOU, and LMP data for the time spanning $t, t + T_H$.[7] The optimization is implemented in Gurobi due to the non-convex binary variable constraints.[8] The resulting returned outputs are stored in an array for later analysis.

4.5 Experiments

To demonstrate MPC, the ESS real-time control is determined for two different weeks (the weeks of July 8 and August 12) with horizons $H = \{1, 2, 6, 12\}$ hours. These are compared to an optimization over the full week given all available data for that week (i.e. solving the optimization once, with $T_H = T_w$), and the cumulative profit (LMP revenue + TOU cost, i.e. Equation 1 for a single time t) are compared.

[7] At the beginning of the week $t = 0$, the stored energy of the ESS is set to $0.5 E_{nom}$.

[8] Full code can be found at the author's GitHub: `https://github.com/kmoy14-stanford/aa-203-final-project`

Algorithm 1: Pseudocode for MPC Algorithm

Input: Arrays for the horizon H starting at time t: load `load_opt`, tariff TOU cost `tariff_opt`, LMP pricing `LMP_opt`, initial stored energy `E_0`

Set $E(0) = E(T_H) =$ `E_0` as in Equation (13) and Equation (14);

for $i = 0, 1, 2, ..., T_H$ **do**

  Set $TOU(i)$ as `tariff_opt`(i);

  Set $l(i)$ as `load_opt`(i);

  Set $LMP(i)$ as `LMP_opt`(i);

  Construct the remaining constraints as in Equations (2) - (12);

Solve the resulting optimization $\max(J)$ with J defined as in Equation (1);

Obtain the resulting immediate LMP revenue `rev_0` $= h[LMP(0)(d^*_{LMP}(0) - c^*_{LMP}(0))]$;

Obtain the resulting immediate TOU reduced cost `cost_0` $= h[TOU(0)(g^*(0))]$;

Obtain the next time step's stored energy `E_1` $= E(1)$;

return *rev_0, cost_0, E_1*

5 Experiment Results

The results of the experiments are summarized in Table 2. In particular, with 1 hour of horizon (corresponding to $1/h = 4$ horizon time steps), the model predictive controller does not have enough foresight to accurately predict the future costs in the system, and performs poorly, resulting in a negative cumulative net profit. The resulting dispatch is also further inhibited by the terminal constraint forced in Equation 14, as the MPC is solving an optimization that constrains the dispatch to return the ESS to the same starting stored energy E_0 within the short horizon. This can be seen in Figure 2b, where the stored energy for 1 hour of horizon has much less variation than for longer horizons.

Table 2: Summary of results. Computations performed on a Macbook Pro (Apple M1 chip, 8-core, 16GB memory).

Week	Metric	Horizon [hr]				
		1	2	6	12	full week
1	Cumulative Net Profit [$]	-1.622	3.991	7.127	7.340	7.521
	Compute Time [min:sec]	00:37	01:14	03:38	07:25	00:07
2	Cumulative Net Profit [$]	-1.922	3.988	5.983	6.321	6.344
	Compute Time [min:sec]	00:37	01:14	03:38	07:25	00:07

In general, increasing the horizon increases the cumulative net profit within the week of operation, approaching the cumulative net profit and stored energy profile obtained by optimizing over the entire week at once. However, this comes at the cost of increasing compute time, where for 12 hour horizon the compute time is half that of the 15-minute interval of the data. This means that the half of the real-time operation of the ESS is spent determining the control for the next time step, for a marginal increase in profit over a 6 hour horizon (which only took 3 minutes 38 seconds to run). Both the 6-hour and 12-hour horizons resulted in near-optimal performance, but if constrained by computational power, the 6-hour horizon should be sufficient for real-time operation.

6 Conclusions and Future Work

A framework to allow the non-simultaneous participation of grid-scale ESS was presented, and the optimization problem for optimal control of the ESS was described. An MPC for real-time control was constructed and demonstrated for one week of dispatch for a residential ESS Tesla Powerwall system, showing the trade-off between MPC horizon and computational power needed for such real-time control. Future work includes substituting in real-time acquired knowledge of the load-TOU-LMP state space (i.e. forecasting) into the model predictive controller and assessing the performance in the same manner as above. Furthermore, the ESS can be treated in more detail using models that account for degradation and capacity loss, and develop control policies that limit the degradation of the ESS while still delivering a net profit for the system.

References

[1] Minjian Cao, Qingshan Xu, Xiaoyang Qin, and Jilin Cai. Battery energy storage sizing based on a model predictive control strategy with operational constraints to smooth the wind power. *International Journal of Electrical Power & Energy Systems*, 115:105471, 2020.

[2] Yankai Cao, Seong Beom Lee, Venkat R. Subramanian, and Victor M. Zavala. Multiscale model predictive control of battery systems for frequency regulation markets using physics-based models. *Journal of Process Control*, 90:46–55, 2020.

[3] Bolong Cheng and Warren B. Powell. Co-optimizing battery storage for the frequency regulation and energy arbitrage using multi-scale dynamic programming. *IEEE Transactions on Smart Grid*, 9(3):1997–2005, 2018.

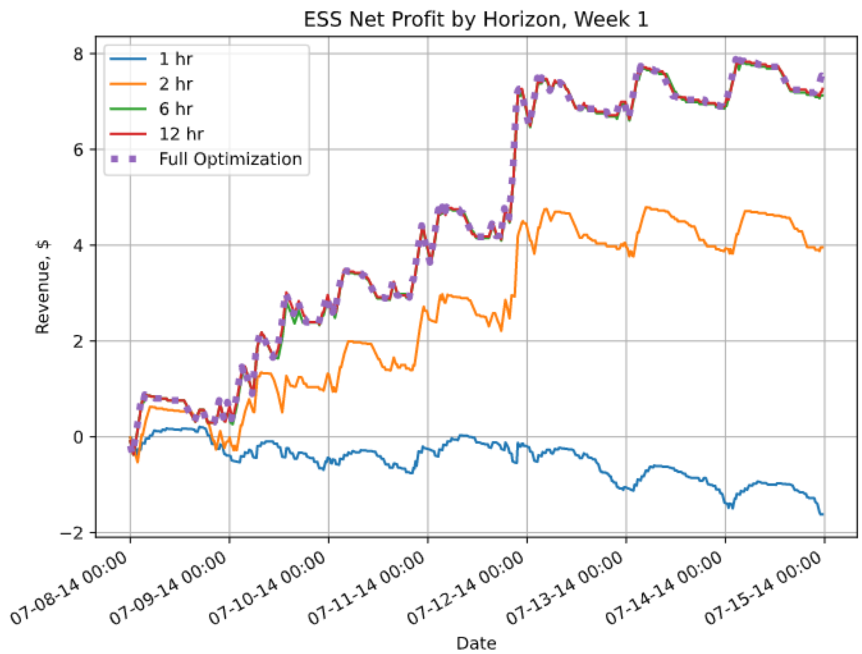
[4] Garett Fitzgerald, James Mandel, Jesse Morris, and Herve Touati. The economics of battery energy storage. Technical report, Rocky Mountain Institute, 9 2015.

[5] Kaitlyn Garifi, Kyri Baker, Dane Christensen, and Behrouz Touri. Control of energy storage in home energy management systems: Non-simultaneous charging and discharging guarantees, 2019.

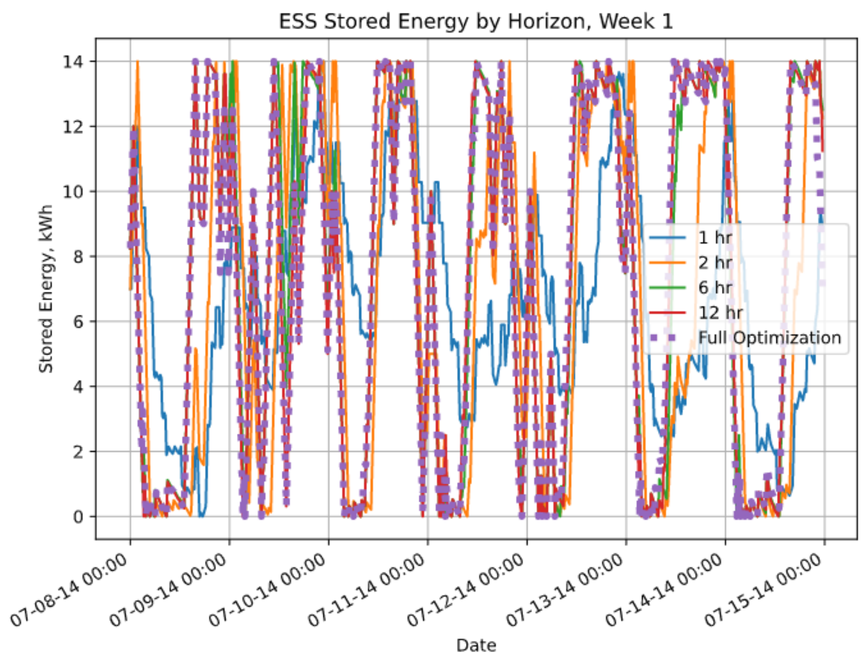
[6] Wai Shin Ho, Sandro Macchietto, Jeng Shiun Lim, Haslenda Hashim, Zarina Ab. Muis, and Wen Hui Liu. Optimal scheduling of energy storage for renewable energy distributed energy generation system. *Renewable and Sustainable Energy Reviews*, 58:1100–1107, 2016.

[7] Guangchun Ruan, Dongqi Wu, Xiangtian Zheng, Haiwang Zhong, Chongqing Kang, Munther A. Dahleh, S. Sivaranjani, and Le Xie. A cross-domain approach to analyzing the short-run impact of covid-19 on the us electricity sector. *Joule*, 4(11):2322–2337, 2020.

[8] Yuanyuan Shi, Bolun Xu, Di Wang, and Baosen Zhang. Using battery storage for peak shaving and frequency regulation: Joint optimization for superlinear gains. *IEEE Transactions on Power Systems*, 33(3):2882–2894, 2018.



(a)



(b)

Figure 2: (a) Cumulative net profits, by horizon, and (b) stored energy profiles as a function of time, by horizon, for Week 1 of TOU-LMP MPC. "Full optimization" denotes an optimization over the entire week at once, with full information of the system load, TOU, and LMP data.