Optimal Operation of Grid-Tied Energy Storage Systems
Considering Detailed Device-Level Battery Models

Zach Taylor, Member, IEEE, Hossein Akhavan-Hejazi, Member, IEEE,
and Hamed Mohsenian-Rad, Senior Member, IEEE

Abstract—The current optimization-based algorithms to operate grid-tied battery energy storage systems (BESS) typically do not look much under the hood of the BESS, i.e. the device-level characteristics of the batteries. This is often due to modeling as well as optimization complexities. However, simplified models may significantly degrade the performance of BESS operation in practice. Therefore, in this paper, we propose a new BESS scheduling optimization framework that accounts for features such as cell-to-cell variations in (a) maximum capacity, (b) charge level balance, and (c) internal resistance. The proposed framework is in the form of tractable mixed integer linear programs. Our approach is to estimate and update the device-tractable models continuously, without the need to interrupt normal operation. We validate the performance compared to an offline approach which is based on dedicated model testing and calibration. To assure accurate performance evaluation, this study also includes developing a power hardware-in-the-loop (PHIL) test-bed that allows for flexible operation and detailed monitoring of BESS under different design scenarios.

Keywords: Optimal BESS operation, peak load shaving, battery pack model, PHIL testbed, cell imbalance, cell capacity.

NOMENCLATURE

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>n, m</td>
<td>Index variables for cell number</td>
</tr>
<tr>
<td>t, τ</td>
<td>Index variables for timeslot</td>
</tr>
<tr>
<td>α, q, u, w</td>
<td>Index variables for discretization</td>
</tr>
<tr>
<td>N</td>
<td>Set of all battery cells</td>
</tr>
<tr>
<td>T</td>
<td>Set of all timeslots</td>
</tr>
<tr>
<td>Q</td>
<td>Set of segments in binary expansion</td>
</tr>
<tr>
<td>π</td>
<td>Number of battery cells in series</td>
</tr>
<tr>
<td>δ, σ, ζ</td>
<td>Length of binary expansion segments</td>
</tr>
<tr>
<td>c₀</td>
<td>Initial stored energy of a battery cell</td>
</tr>
<tr>
<td>r</td>
<td>Internal resistance of a battery cell</td>
</tr>
<tr>
<td>P₀</td>
<td>Output power of a battery cell</td>
</tr>
<tr>
<td>Pᵢₙ</td>
<td>Internal power generated/absorbed by a cell</td>
</tr>
<tr>
<td>Vᵦ₀</td>
<td>Open circuit voltage of a battery cell</td>
</tr>
<tr>
<td>V₀</td>
<td>Terminal voltage of a battery cell</td>
</tr>
<tr>
<td>I</td>
<td>Current into a battery cell</td>
</tr>
<tr>
<td>C</td>
<td>Stored energy of a battery cell in Wh</td>
</tr>
<tr>
<td>SoC</td>
<td>State of charge of a battery cell, in percent of the cell energy capacity in Wh</td>
</tr>
<tr>
<td>C</td>
<td>Maximum cell energy in Wh</td>
</tr>
<tr>
<td>τ, i</td>
<td>Maximum cell voltage and current limits</td>
</tr>
<tr>
<td>v₀, i</td>
<td>Minimum cell voltage and current limits</td>
</tr>
<tr>
<td>f(·)</td>
<td>Function to relate Vᵦ₀ and C.</td>
</tr>
<tr>
<td>a</td>
<td>Slope of linear function of f(·)</td>
</tr>
<tr>
<td>l</td>
<td>Look up table for f(·)</td>
</tr>
<tr>
<td>B</td>
<td>Binary variable to discretize cells’ current</td>
</tr>
<tr>
<td>D</td>
<td>Binary variable to discretize cells’ Vᵦ₀</td>
</tr>
<tr>
<td>F</td>
<td>Binary variable to discretize cells’ V₀</td>
</tr>
<tr>
<td>φ</td>
<td>Variable to relate V₀ and discrete current</td>
</tr>
<tr>
<td>ψ</td>
<td>Variable to relate Vᵦ₀ and discrete current</td>
</tr>
<tr>
<td>Ω</td>
<td>Variable to relate discrete Vᵦ₀ and current</td>
</tr>
<tr>
<td>Γ</td>
<td>Variable to relate discrete V₀ and current</td>
</tr>
<tr>
<td>L</td>
<td>Large number for the Big L binary method</td>
</tr>
<tr>
<td>α, β, γ</td>
<td>Discretization parameters</td>
</tr>
</tbody>
</table>

A. Background and Motivation

Stationary battery energy storage systems (BESS) are gradually adopted in practice for various grid applications. A common approach in the existing algorithms that tend to “optimally” schedule BESS for grid integration is to use simplistic models of the battery components to make it mathematically tractable to obtain an optimal solution. The battery is often assumed as one large single battery cell with known maximum energy capacity, that for the most part is ideal, except for having a constant round-trip efficiency of charge/discharge. The remaining usable battery energy is therefore calculated based on the net sum of energy that is put into or taken from the battery, while taking into account the round trip efficiency, c.f. [1]–[7]. As a result, it is often overlooked that the maximum capacity of a battery may vary from its nominal ratings, in particular because it is often a pack of many cells, that are combined in series and/or in parallel. Aspects such as imbalance in the cells charge level, variable maximum capacity across different cells, and variable losses at different power rates and charge levels, all affect the BESS operation [8].

The properties listed above, may create operational issues when a simplified model of a battery pack is used in practice. These issues were encountered in practice, for example in the study in [9] for the case of a grid-connected 100 kW / 500 kWh battery testbed, as shown in Fig. 1. At Event 1, the battery pack unexpectedly sets its output power to zero during discharge. This was caused by a single cell within the pack reaching its minimum voltage, triggering a halt signal from the Battery Management System (BMS), reporting zero remaining usable charge for the entire pack. Whether this cell was limited by being out of balance or low cell energy capacity, the algorithm in [9] did not account for this behavior. Such unexpected halts in BESS operation may occur in critical times and result in significant loss of performance of the BESS application.
B. Summary of Contributions

Motivated by the above and other similar practical examples in the literature, in this paper, we investigate whether accounting for the battery characteristics at cell-level can enable us to better optimize the charge/discharge schedule for the battery pack and prevent unexpected halt events. The main contributions in this paper can be summarized as follows:

- An extended BESS scheduling optimization problem is formulated that incorporates the device-level model of the BESS that accounts for the battery pack construct, cell imbalance, and other cell characteristics. Such device-level details are often not used in the comparable literature when it comes to the optimization problems to operate grid-tied BESS, because of their added complexity, namely creating non-convexity in the problem formulation, as well as imposing challenges to estimate the parameters of the model during the BESS operation.

  - We address the first challenge, i.e., with respect to non-convexity, by proposing two rigorous mathematical approaches to reformulate the constraints. The first approach, referred to as CALVS, combines a linear approximation of voltage versus state-of-charge curve with a binary expansion of bilinear terms involving current. The second approach, referred to as CANVS, rather applies binary expansion to the bilinear terms involving voltage; and then combines it with look-up table approximation of the voltage versus state-of-charge curve. Both methods significantly reduce computational complexity to solve the BESS scheduling optimization problem.

  - We address the second challenge, i.e., the need for on-the-go estimation of model parameters, by developing an event-based calibration method, making the parameter estimation non-intrusive, which allows the battery to run continuously while learning the best way to operate.

C. Literature Review

This paper is not about developing new battery cell models; it is rather about adopting an adequate existing battery cell model into the problem formulation of optimization-based grid-tied BESS operation. Looking at the models available, the choices are broad. In fact, a wide range of different empirical, c.f. [10], [11], circuit-based, c.f. [12]–[14], and physics-based, c.f. [15], [16], battery models and modeling approaches, as well as hybrids of such approaches, c.f. [17], have been developed in the literature. A summary of these models are also provided in review papers, such as [18]–[20], and the references therein.

Broadly speaking, we may categorize the existing models based on how they can be used in applications such as (i) battery design, simulation, and characterization, c.f. [21], [22] or hardware-embedded environment, e.g. in [23]; (ii) battery characterization, parameter estimation, and/or state estimation/prediction in BMS, c.f. [24]; and (iii) real time control, e.g. in [25]–[27], or extended-time control and operation management, e.g. in [2]–[4], [9], [28], [29].

Unlike the models in groups (i) and (ii), the models used for control of the BESS, i.e., in group (iii), often do not respond to a given power output, rather need to search for a feasible path in all operation scenarios in a short or extended time horizon. Accordingly, the challenges of incorporating complex models are much higher. For real-time SoC management of the cells, and SoC balance across the pack, several simpler circuit-based models in feedback control loops, or extremum seeking models are presented in [11], [25]–[27], [30]. These either search for feasible/stable solution and/or sub-optimal solution.

In contrast, for extended time operation and management applications, the complexity of non-linear, non-convex battery
models is significantly more challenging. Thus, the majority of models either use ideal storage with simple Coulomb counting, or introduce charge/discharge efficiency factor and certain complementarity constrains, which again leads to MILP problems, c.f. [2]–[4], [4], [9]. There are efforts to add more details of the batteries to these basic models, while preserving the convexity and/or tractability, e.g. by adding simple empirical terms to account for degradation [4], [28], [31], [32], charge rate [3], [28], temperature [3], [28], and depth of discharge [3], [28]. There are also studies that use non-linear models using the gradient seeking solvers, without any regards for convexity or sub-optimality, e.g. in [33]. Recent works try to address the non-convexity by using certain convex-relaxation techniques, [34], [35] and/or conditions that the complementarity constraints can be dropped [36]. Thus, there still remains a gap in literature for optimization-BESS operation to use models and estimation approaches that consider battery system features and are able to track cell-level variations over time.

In addition, the common practice in adopting battery models, particularly for BESS operation optimization, is to treat the entire BESS as a large battery cell, rather than to account for the impact of the underlying battery pack construct and conditions [37]. Accordingly, often the models of a single cell are simply scaled-up to utility size BESS by applying scaling factors to the maximum and minimum capacities of the cell, e.g., in [1], [38], [39]. This is done by an implicit assumption of uniformity across the cells which is often the case in newly deployed systems. However, this can be problematic at times, especially for the aged battery packs, and thus for BESSs in their later portion of their operation life. In a large utility-scale battery pack, it is common to have many series- and parallel-connected battery cells combined to create a high maximum capacity as well as high voltage battery pack [37]. The management of such interconnected battery cells is often done through a BMS by balancing the voltage of each cell [28] and may have a noticeable impact on operation performance of BESS, but is not factored into these scaled up models.

II. PROBLEM STATEMENT: BATTERY PACK IMBALANCE MODELING FOR BESS OPTIMIZATION

During BESS operation in practice, a BMS often monitors the voltage of individual cells. The energy availability of a battery pack is often determined by the BMS based on the voltage at the battery cell terminals [40]. If the voltage of any battery cell within a battery pack falls below or exceeds above certain thresholds, then the whole battery pack is considered to be depleted or fully charged, respectively. Violating these voltage limits for any battery cell will damage the cell and must be avoided. This is enforced by the BMS.

However, not all cells within a battery pack will have the same voltage trajectory during charge/discharge. For instance, consider the voltage profiles of a lab-scale BESS, shown in Fig. 3, during one cycle of discharge and charge. While the 12 cells in this battery are discharged and charged with the same current, the voltage of some cells decline and incline faster than others. Additionally, when the pack begins to charge, some of the cells may already have some energy stored. Thus, in each cycle, the available energy of a battery is often affected by cells’ initial energy level and maximum capacity variations. The BMS tracks these variations in realtime, applies charge balancing, and estimates the current SoC from cell voltage measurements. However, neither BMS nor the current typical BESS schedule algorithms produce and account for estimates of the battery pack states in future operating schedules.

A. Circuit Representation of Battery Pack in Optimization

The power drawn/injected in the battery pack at any time slot $t \in T$, denoted by $P_{BESS}[t]$, is the sum of the power drawn/injected in each individual cell $n$, denoted by $P_o[t, n]$:

$$P_{BESS}[t] = \sum_{n \in N} P_o[t, n]. \quad (1)$$

Here $P_o[t, n]$ can take both positive and negative values, where negative values indicate discharging of the cell.

Each cell in a battery pack can be represented by a circuit as shown in Fig. 4, with a variable voltage source and a resistance $R$ as parameters. Note that, $V_{oc}$ is a non-linear non-convex function of the $SoC$, see the equation in (4) and the curve in Fig. 5.

![Fig. 3. The voltage profiles of the cells in a used battery pack.](image)

![Fig. 4. Circuit-based cell model with $V_{oc}$ as a function of SoC and internal resistance $R$ as parameters. Note that, $V_{oc}$ is a non-linear non-convex function of the $SoC$, see the equation in (4) and the curve in Fig. 5.](image)
where the notations are defined in the nomenclature. All cells are in series and their current is the same. Importantly, the open circuit voltage $V_{oc}$ is a function of $SoC$. We define $SoC$, as a normalized measure of a cell’s stored energy, defined as the ratio of the stored energy over the total energy able to be stored\(^1\). We denote $C$ as the amount of the energy in Wh, which is stored in a cell when fully charged from a depleted state. Accordingly, the amount of energy in Wh stored in cell $n$, $C[t; n]$, when it is at $SoC[t; n]$ is expressed by $SoCC$.

Both $V_{oc}$ and $V_{o}$ are positive and typically take values between 2.5 and 4.0 \([41]\) . The above system of equations provides a discrete state-space representation for the battery cell output voltage $V_{o}$ in terms of input current $I$; and their product in (2) indicates the output power $P_{o}$.

Additionally, the BMS enforces upper and lower limits on the terminal voltage of each battery, denoted by, $\bar{v}_{o}$ and $v_{oc}$. This can be expressed by the following inequality constraints that must hold for all $t \in T$ and all $n \in N$:

$$v_{oc} \leq V_{o}[t; n] \leq \bar{v}_{o}. \quad (8)$$

It is worth emphasizing that, from (2) and (3), since $V_{oc}$ is a function of $SoC$, both $V_{o}$ and $P_{o}$ are also functions of $SoC$. As we will discuss in Section III, the relationship between $V_{oc}$ and $SoC$ is nonlinear and non-convex. Therefore, the system of equalities and inequalities in (1)-(8) is nonlinear and non-convex, i.e., hard to be used in an optimization problem.

One may ask whether the above circuit-based model is sufficiently complex and accurate. The answer would depend on which models and for which applications one is comparing the above model with. If we compare the above model with the typical models that are used in the literature on optimization-based grid-tied battery operations, then the answer is ‘yes’, because this model is significantly more complex than the simple models that have been used previously, e.g., see [2]–[4], [9], [28], [31], [32]. But if we compare this model with the models that are used in the literature, for example, on cell-level battery simulation, then the answer is ‘no’ because battery modeling on its own is a rich field, e.g., see [12]–[14].

But the latter would not be a fair comparison, because this paper is about optimization-based grid-tied BESS operations. The focus in this paper is not on developing any new battery model; it is rather on tackling the challenges in bringing the above existing cell-level battery model into the context of optimization problem formulation of grid-tied BESS operation, and also to explain and showcase the insights and advantages that one can gain by doing so.

We will study the impact of battery model accuracy on the performance of the peak shaving problem in Section V.E.

Last but not least, it should be noted that the impact of temperature variation is not included in this model, though there exist circuit models to account for temperature, see [16]. Since we focus specifically on stationary grid-tied BESS in this paper, we make the practical assumption that the battery is housed in a container or building and the temperature is controlled and held constant via an HVAC system. Thus we ignore the impact of temperature on the battery operation.

\(^1\)Accordingly, the term state-of-energy (SoE) can also be used here.

---

**B. BESS Scheduling Optimization Problem**

Equations (1)-(8) provide a mathematical representation of each battery cell, which can be used to identify the device-level operational constraints in any BESS scheduling optimization problem for grid-integration purposes. Different objective functions can be considered. Without loss of generality, we consider the problem of peak load shaving, which is one of the prominent family of grid-tied applications of BESS \([42]–[44]\). The objective is to minimize the net peak load of a building \([42], [44]\), a distribution feeder \([9]\), or a substation \([45]\). Mathematically, this problem can be formulated as

$$\begin{align*}
\text{minimize} & \quad \max_{t} (P_{load}[t] + P_{BESS}[t]) \\
\text{subject to} & \quad \text{Eqs. (1)-(7)}.
\end{align*} \quad (9)$$

where $P_{load}[t]$ denotes the background load at time slot $t$ in the target load shaving problem, such as the load of the building, feeder, or substations, without the impact of the BESS. Recall that if $P_{BESS}[t]$ is positive then the BESS is being charged; and if $P_{BESS}[t]$ is negative then the BESS is being discharged. Sometimes, one may also add a regularization term $\|P_{BESS}[t]\|$ to the objective function to minimize the use of the BESS, c.f. \([9]\). Similarly, additional constraints can also be considered such as policy constraints \([46]\), financial constraints \([35], [47]\), or other non-BESS related physical constraints \([48]\).

Our focus in this paper is specifically on the constraints that capture the essence of the BESS system, i.e., equations (1)-(8). Importantly, such constraints are often not used in the literature on optimization-based operation of a grid-tied BESS. In fact, none of the references that we cited in the previous paragraph, i.e., \([35], [46]–[48]\) , consider the constraints of the type in (1)-(7) . They use rather simplified battery models that do not take into account imbalance in the cell energy levels, variable maximum capacity across different cells, and variable losses at different power rates and charge levels. While we acknowledge the fact that such detailed battery cell-level models are often not used in the literature in optimization-based operation of a grid-tied BESS, in this paper, we argue that, we think we should look under the hood of the BESS, i.e. the device-level characteristics of the batteries, in order to improve the efficiency and reliability of the grid-tied BESS operation. We
also acknowledge the fact that one of the main reasons not to include such detailed cell-level models in this context is the added complexity to the problem. Specifically, regardless of the objective function being used, the optimization problem in (9) is difficult to work with because of two major challenges:

- **Challenge 1: Non-Convex Formulation:** There are two sources of non-convexity in (1)-(8). First, \( V_{oc}[t, n] \) is a non-convex function \( f() \) of \( SoC[t, n] \) in (4). This is shown in 5. Second, the constraints in (2) and (7) involve bi-linear terms, which are also non-convex [48].

- **Challenge 2: Parameter Estimation During Operation:** The difficulty in estimating the parameters that exist in the above circuit model of the battery cells, unless we disintegrate them from the grid and test them individually.

The main contribution in this paper is to address the above two challenges and pave the way to allow the tractable linearization of the bilinear terms in (2) and (7). Suppose the usable remaining potential, the lowest open cell voltage, is a nonlinear sigmoid-shaped function of \( I \). Each segment is denoted by an index \( t,n \). Therefore, one may choose to approximate function \( f() \) with a single line, as shown in Fig. 5(a):

\[
V_{oc}[t, n] = v_{oc} + a C[t, n].
\]

Here \( a \) and \( v_{oc} \) denote the slope and vertical intercept for the linear function under linear approximation. In this approximation, the lowest open cell voltage \( v_{oc} \) is associated with zero remaining usable battery energy, i.e., \( C[t, n] = 0 \). Thus, the linear \( SoC \) curve is only fit in the region from 20% to 90%.

2) **Binary Expansion of Current \( I[t] \):** Since the current that goes through the battery pack is always bounded in practice, we can apply the binary expansion method, c.f. [48], to linearize the bilinear terms in (2) and (7). Suppose the operation range of current is divided to \( q \) segments of equal length \( \delta = (\bar{I} - \underline{I})/q \). Each segment is denoted by an index \( q \), where \( q \in Q = [1, \ldots, q] \). Let \( \beta[q] = \bar{I} + \delta q \) denote the amount of current at segment \( q \). At each time slot \( t \in T \), we define binary variables \( B[t, q] \) to approximate \( I[t] \) as follows:

\[
I[t] - \delta/2 \leq \sum_{q \in Q} \beta[q] B[t, q] \leq I[t] + \delta/2 \tag{11}
\]

Here, the goal is to approximate variable \( I[t] \) with the closest value in the constant vector \( \beta[q] \). The selection is done via binary variable \( B[t, q] \). From (12), for each time slot \( t \), the term \( B[t, q] \) must be 1 for exactly one choice of \( q \), where \( q = 1, \ldots, q \). From this, together with (11), the summation

\[
\sum_{q \in Q} \beta[q] B[t, q] \tag{13}
\]

is always equal to \( \beta[q] \) for exactly one choice of \( q \), for which \( B[t, q] = 1 \). The whole purpose here is to quantize \( I[t] \) such that we can use techniques from binary optimization to transform the non-convex bilinear terms that involve \( I[t] \) into tractable linear terms, as we will see in the next paragraph.

The approximation error is bounded by \( \delta/2 \). Thus, we can arbitrarily improve the approximation accuracy by decreasing \( \delta \) or equivalently increasing \( q \). Of course, this will be at the expense of increasing the computational complexity of the model because of the increasing number of binary variables.

Next, for each time slot \( t \) and each cell \( n \), we introduce auxiliary variables \( \Psi[t, n, q] \) and \( \Phi[t, n, q] \) in relationship with \( V_{oc}[t, n] \) and \( V_{o}[t, n] \), respectively, as shown below:

\[
0 \leq V_{oc}[t, n] - \Psi[t, n, q] \leq (1 - B[t, q])L \tag{14}
\]

\[
0 \leq \Psi[t, n, q] \leq B[t, q]L \tag{15}
\]

\[
0 \leq V_{o}[t, n] - \Phi[t, n, q] \leq (1 - B[t, q])L \tag{16}
\]

\[
0 \leq \Phi[t, n, q] \leq B[t, q]L \tag{17}
\]

Here \( L \) is a large number. From (14) and (15), \( \Psi[t, n, q] \) is equal to \( V_{oc}[t, n] \) at any \( q \) for which \( B[t, q] = 1 \); and zero otherwise. Similarly, from (16) and (17), \( \Phi[t, n, q] \) is equal to \( V_{o}[t, n] \) at any \( q \) for which \( B[t, q] = 1 \); and zero otherwise.

We have introduced these constraints so that we have auxiliary variables to approximate the bilinear equality constraints in (2) and (7) as the following linear constraints:

\[
P_{o}[t, n] = \sum_{q \in Q} \beta[q] \Psi[t, n, q], \tag{18}
\]

\[
P_{o}[t, n] = \sum_{q \in Q} \beta[q] \Phi[t, n, q]. \tag{19}
\]

3) **Optimization Problem under Approach 1:** We are now ready to reformulate the optimization problem in (9) as

\[
\begin{align*}
\text{minimize} & \quad \max (P_{\text{load}}[t] + P_{\text{BESS}}[t]) \\
\text{subject to} & \quad \text{Eqs. (1), (3), (5), (7), (10)} \\
& \quad \text{Eqs. (9) – (17)}.
\end{align*}
\]

The above problem is a mixed-integer linear program (MILP).
B. Approach 2: CANVS

Since the cell voltages change rapidly at the SoC extremes, the linear model under Approach 1 may not fit the battery characteristics well around the low cell charge levels. Thus, we propose a second approach, CANVS, to further improve the accuracy of battery model by using a look up table (LUT) of the values for $V_{oc}$ at various charge levels. This improvement, however, comes at the cost of a higher quantity of binary variables and thus higher computational requirements.

1) Binary Expansion of Voltage $V_{oc}[t,n]$:

In this second approach, we apply the binary expansion method to $V_{oc}[t,n]$ instead of to $I[t]$. First, we divide the operation range of open-circuit voltage for each cell into $\pi$ sections of equal length $\sigma = (\bar{V}_{oc} - \underline{V}_{oc})/\bar{u}$. Each segment is denoted by an index $u$, where $u \in U = [1, \ldots, \pi]$. Let $\alpha[u] = \underline{V}_{oc} + \sigma u$ denote the amount of voltage at segment $u$. At each time slot $t \in T$ and for each battery cell $n$, we define binary variables $D[t,n,u]$ in order to approximate $V_{oc}[t,n]$ as follows:

$$V_{oc}[t,n] - \sigma/2 \leq \sum_{u \in U} \alpha[u] D[t,n,u] \leq V_{oc}[t,n] + \sigma/2 \quad (21)$$

and

$$\sum_{u \in U} D[t,n,u] = 1. \quad (22)$$

Here, the goal is to approximate variable $V_{oc}[t,n]$ with the closest value in the constant vector $\alpha[u]$. The selection is done via binary variable $D[t,n,u]$. From (22), for each time slot $t$ and each cell $n$, the term $D[t,n,u]$ must be 1 for exactly one choice of $u$, where $u = 1, \ldots, \bar{u}$. From this, together with (21), the following summation

$$\sum_{u \in U} \alpha[u] D[t,n,u] \quad (23)$$

is always equal to $\alpha[u]$ for exactly one choice of $u$, for which $D[t,n,u] = 1$. The whole purpose here is to quantize $V_{oc}[t,n]$ such that we can use techniques from binary optimization to transform the non-convex bilinear terms that involve $V_{oc}[t]$ into mathematically tractable linear terms.

2) Look-Up Table Approximation of Function $f(\cdot)$:

Next, consider the nonlinear curve $f(\cdot)$ in Fig. 5 for each cell $n$. Each cell may have a unique curve. Suppose $l[n,u]$ denotes cell $n$’s charge at voltage level $\alpha[u]$. Accordingly, at each time slot $t \in T$ and for each cell $n$, we can write:

$$C[t,n] = \sum_{u \in U} l[n,u] D[t,n,u]. \quad (24)$$

Here, the pairs $\alpha[u]$ and $l[n,u]$ form a look-up table approximation of the nonlinear curve $f(\cdot)$ in Fig. 5 for each cell $n$. The approximation accuracy in (21) can be arbitrarily improved by decreasing $\sigma$ or equivalently increasing $\bar{u}$.

Next, we tackle the bilinear terms in (2) and (7). Similar to the setup in (14) and (15) we define a new auxiliary variable $\Omega[t,n,u]$ in relationship with $I[t]$ for each time slot $t \in T$ and each cell $n \in N$, through the following constraints:

$$(D[t,n,u] - 1)L \leq I[t] - \Omega[t,n,u] \leq (1 - D[t,n,u])L \quad (25)$$

We can now approximate the bilinear constraint in (7) as:

$$P_{in}[t,n] = \sum_{u \in U} \alpha[u] \Omega[t,n,u]. \quad (27)$$

As for the bilinear constraint in (2), we also apply binary expansion to $V_o[t,n]$. The methodology is identical to how we did this for $V_{oc}[t,n]$. We just need to replace $u, \bar{u}, \sigma, \alpha[u], U,$ and $\bar{D}[t,n,u]$ with $w, \bar{w}, \zeta, \gamma[u], W,$ and $\bar{F}[t,n,w]$, respectively. The corresponding new constraints become:

$$V_o[t,n] - \zeta/2 \leq \sum_{w \in W} \gamma[w] F[t,n,w] \leq V_o[t,n] + \zeta/2 \quad (28)$$

$$\sum_{w \in W} F[t,n,w] = 1. \quad (29)$$

Here, the goal is to approximate variable $V_o[t,n]$ with the closest value in the constant vector $\gamma[w]$. The selection is done via binary variable $F[t,n,w]$. From (29), for each time slot $t$ and each cell $n$, the term $F[t,n,w]$ must be 1 for exactly one choice of $w$, where $w = 1, \ldots, \bar{w}$. From this, together with (28), the following summation

$$\sum_{w \in W} \gamma[w] F[t,n,w] \quad (30)$$

is always equal to $\alpha[u]$ for exactly one choice of $w$, for which $F[t,n,w] = 1$. The whole purpose here is to quantize $V_o[t,n]$ such that we can use techniques from binary optimization to transform the non-convex bilinear terms that involve $V_o[t]$ into tractable linear terms, as we will see in the next paragraph.

We can now replace constraint (2) with the following:

$$(F[t,n,w] - 1)L \leq I[t] - \Gamma[t,n,w] \leq (1 - F[t,n,w])L \quad (31)$$

$$-F[t,n,w]L \leq \Gamma[t,n,w] \leq F[t,n,w]L, \quad (32)$$

$$P_{in}[t,n] = \sum_{w \in W} \gamma[w] \Gamma[t,n,w] \quad (33)$$

3) Optimization Problem under Approach 2: The MILP formulation under this approach is obtained as follows:

$$\text{minimize} \quad P_{o,B,C,P_o,t,I,V_o,V_m,T,\Omega,D,F} \quad \text{max}(P_{load}[t] + P_{BESS}[t])$$

subject to \ Eqs. (1), (3), (5), (7), (19) – (29).

IV. CHALLENGE 2: PARAMETER ESTIMATION

A. Battery Operation Tests with PHIL Testbed

The testbed is shown in Fig. 2. It includes all the key components of a typical grid-tied BESS, implemented via hardware devices, RTDS models, and prediction/scheduling programs and data-logging. The testbed contains 12 Lithium Iron Phosphate battery cells each at 40 Ah, and 3.3 V nominal voltage [41]. The cells are connected in series forming a 40 Ah, 39.6 V battery array. The batteries can charge/discharge at currents as high as ±10C rate, but is limited in our test for safety and to extended lifetime. The battery pack is in the used condition; thus, the maximum capacity of some cells are
reduced. The cells are also at different levels of charge, i.e. the pack is unbalanced. Additional details about the architecture, hardware components, software systems, communications, and computer controls of this PHIL testbed are explained in [49].

We performed three types of battery operation tests using the BESS PHIL system in order to train the battery models:

**Test #1:** In this test, the BESS is operated for 120 hours which consists of multiple charge and discharge cycles across its entire allowed range of operation, i.e. from 20% to 90% of SoC (as reported by the BMS), using a low constant current as shown in Fig. 6(a). We charged/discharged the battery pack at the constant current rates of 1 A and 2 A. Since the battery is not discharged below C = 20% or above τ = 90% of its capacity, we can assume that it reaches its full usable capacity, or fully charged, at 90% SoC; and to its zero usable capacity, or fully depleted, at 20% SoC. This scaling task does not impact any results, yet it simplifies explaining the results.

**Test #2:** This test involves a long idle time (i.e. time with no charge or discharge current, I[t] = 0) immediately after a charge cycle, as shown in Fig. 6(a). We monitored the battery pack while it was idle for 80 hours right after it was charged at constant current of 2 A, allowing the cells’ terminal voltages to settle to their Voc values. We use the test results later in this section to estimate cell internal resistance.

**Test #3:** This test operates the PHIL testbed for 7 days based on a typical daily operation schedule of a BESS. We use this test to create data for a very realistic scenario based on real-world Time-Of-Use (ToU) pricing, as shown in Fig. 6(b). We validate the models ultimately based on the data from this test.

### B. Parameter Estimation: Internal Cell Resistance

This can be done by using the measurements in Test #2 and based on the model depicted in Fig. 4. The estimation process is the same for both optimization-based approaches. If we stop charging (or discharging) the battery, but we continue monitoring the terminal voltages, then we can see that, the terminal voltage gradually converges to Voc at that specific charge level. The cell’s resistance can therefore be obtained from (3) once we calculate the values of Voc, Voc, and I.

The above offline estimation method may not be desirable in practice; because it requires the BESS to be idle for a long time. This prevents the BESS from normal operation. Alternatively, for online parameter estimation, the value of r[n] can be obtained by looking at the data from moments before and moments after a change in current. Recall from Section II, that we model each battery cell as a voltage source Voc[t, n] in series with a constant resistance r[n]. The value of Voc[t, n] is governed by C[t, n]. Therefore, from (3), if we take two operating points t1 and t2 close together in time, but with a measurable difference in current, then we can assume that C[t1, n] − C[t2, n] is equal to zero, implying that the open circuit voltage Voc[t, n] is constant. We can then solve for r[n]:

$$r[n] = (V_o[t_2, n] - V_o[t_1, n]) / (I[t_2] - I[t_1]).$$  \hspace{1cm} (35)

Fig. 7 shows the results of estimating cell resistance using the aforementioned methods. The results are somewhat (but not significantly) different across the two methods. The difference in the estimated resistance is consistent among the cells. This means that the online approach leads to lower estimated values for the resistance; which can be offset, if need be. Also, since the overall modeling accuracy is measured rather based on how closely the complete model can follow the measured data, we do have some tolerance in the resistance estimation inaccuracy, as it is compensated by the process of estimating the Voc curve. Thus, we can conclude that the online method is a reasonably accurate practical method, when compared with the offline method that is intended for model validation.
C. Parameter Estimation: Function $f(\cdot)$ – Linear Model

One option that is suggested in the literature, e.g. in [50], is to charge and discharge the battery at low current rates, assuming $V_{oc}[t, n]$ is equal to the terminal voltage. However, the value of $V_{oc}[t, n]$ varies across charge and discharge cycles, and from one cycle to another. Therefore, in this work, we rather charge and discharge the batteries at 2.5% and 5% C-rate. Even for such small currents, the change of current attains different values for $V_{oc}[t, n]$. We use linear regression, to fit a line to the measured data of different cycles.

We leverage the knowledge of the cells’ maximum voltages and minimum voltages, associated with fully charged and fully depleted cells, given in the manufacturer data sheet. Thus, the linear model involves estimation of only one parameter. That is, for each cell, we only need to estimate the energy associated with the maximum cell voltage. Based on the recorded measurements, we know all the variables and parameters in (10) except for the line slope $a$. For each measured terminal voltage and current, we obtain the measured value for $V_{oc}[t, n]$ from (3), given the estimated value of $r[n]$. Also, we obtain $C[t, n]$ from (6), where $P_{in}[t, n]$ is obtained from (7). Accordingly, we solve a linear regression for the value of $a$ that minimizes the error in (10) between the measured and modeled data.

The results are shown in Figs. 8(a) and (b), for one cell, and for several cells, respectively. We can see that, in addition to different resistance values, different cells also have different $V_{oc}$ curves, i.e., each cell can have varying maximum capacity.

D. Parameter Estimation: Function $f(\cdot)$ – Lookup Table

Here we estimate the curve of $V_{oc}[t, n]$ versus $C[t, n]$ when it is split into $T$ sections. As before, we assume to know the starting point and some information about the shape of the SoC curve from the manufacturer data sheet. We use shape prescriptive curve fitting, c.f. [51], in order to solve the parameter estimation problem. This method uses least squares fitting of splines curves on sample data, subject to constraints such as monotonicity, curvature, and value assignments.

Specifically, we consider the following constraints: (i) The cell open circuit voltage is always increasing with charge level; (ii) the curve must pass through certain points, such as the full and empty voltage limits given by the cell manufacturer; (iii) the minimum open circuit value occurs at zero available energy; (iv) the number and size of sections are predetermined.

Fig. 9 shows the results of the estimated curve fitted to the measured data with the given number of knots. From the estimated curve we can identify arbitrary number of $V_{oc}$ versus $C$ pairs for the look-up table, including the specified knots in the figure. We observed that under a typical daily charge and discharge schedule, the model achieves a better accuracy if we use only a limited duration of historical data for training, thus, we reset the model parameters based on the newest data every time the battery reaches to a calibration event, i.e. when a cell reaches to the lowest allowed voltage.

V. EXPERIMENTAL RESULTS, MODEL VALIDATIONS, PERFORMANCE ASSESSMENT

A. Performance Evaluation Metric and Baseline Model

It is a challenge in practice to compare the accuracy of different battery models, in estimating the charge levels of battery cells. The reason is that the “true” charge level of a battery cannot be directly measured. In fact, even the SoC values reported by the BMS devices, which are sometimes mistaken as true charge level, do contain some error. The only exception is when the battery is depleted, i.e., it reaches its minimum operation voltage. Only at this point, one can be certain that the available energy is practically zero, i.e., the battery has no more usable energy available for discharge.

This is exactly the same moment that we discussed as Event 1 in Fig. 1 in the Introduction. This is also the same moment when a typical BMS device conducts recalibration. We refer to such events as “Zero Charge Level Calibration Events”, or simply “Calibration Events” for ease of discussion.

Consider the estimated energy level at the moment right before a Calibration Event. Clearly, if the model is perfect, then the estimated charge level would be zero at this moment. A higher mismatch for the estimated energy level however means a less accurate model. This mismatch can be used as a metric to assess the accuracy of different battery models.

A baseline model is considered for comparison. This baseline model is referred to as Coulomb Counting Approach.

---

Fig. 8. The results of $V_{oc} - C$ relation estimation based on the linear model. (a) the linear model regression to measured data for the cell Num. 8 after two charge/discharge cycles, (b) the linear models obtained for the twelve cells.

Fig. 9. Curve fitting used in the LUT method, each division bounded by the dotted green line represents one piecewise linear portion of the curve.
(CCA), which is based on summing up the energy delivered at the battery terminals in Ah or percent of rated capacity. This model assumes that the stored energy in the battery pack is the sum of the energy in the pack at the beginning of a time interval plus the energy delivered at the battery terminals during that time interval; that is:

\[ C[t] = C[t-1] + P_{in}[t-1]\eta_{in} + P_{out}[t-1]\eta_{out} \]  

(36)

where \(\eta_{in}\) and \(\eta_{out}\) are the charge and discharge efficiency estimates, respectively.

A modified version of the CCA model is used in practice in BMS devices for real-time estimation of SoC, with the difference that the BMS device periodically calibrates its estimations based on the measured voltage of a battery cell that reaches to the highest or the lowest allowed voltage thresholds. As a result, the estimated charge level periodically jumps up to the nominal capacity and down to zero based on such periodic measured voltages. The calibrations limit the accumulation of the model error over time, but introduce non-continuous SoC estimates. As a result, such models cannot be used for scheduling purposes. Moreover, if the battery pack rated capacity is not accurate, then the calibration will not reach the correct charge level. The algorithm used in BMS devices can be found in Appendix A.

B. Battery Model Accuracy Evaluation Results

The proposed CALVS and CANVS models, along with the baseline CCA model, are trained using data from previous cycles. We then compare the charge level mismatch at the calibration events for each model. Two types of charge and discharge cycles are considered, as we explain next.

1) Constant Current Charge and Discharge Cycles: Here, the evaluation is done based on Test #1. The results on model mismatch error are shown in Fig. 10 and Fig. 11(a). Over the course of the 120 hours of this test, the BESS encountered seven distinct calibration events. When examining the trends in Fig. 10, we can see that CCA has significantly higher error, which accumulates over time. This is due to the integral nature of Coulomb counting used in the model, accumulating error over time. Note that even if the maximum capacity parameter is periodically updated in the CCA model, it cannot differentiate between the capacity change due to the battery pack imbalance versus a change due to degradation or cell variation. The proposed CALVS and CANVS models, however, perform significantly better. They allow us to re-estimate the model parameters and account for the decrease or the increase in pack capacity in successive cycles. The advantage is that they account for changes in the maximum capacity of the cells, and the changes in balance of the pack. It is worth pointing out that the CANVS method appears to perform better than the CALVS method, also see Section V-D.

2) Typical BESS Daily Operation Cycles: This evaluation is done based on Test #3. The results on model mismatch error are shown in Fig. 11(b) for 200 hours of operation. During the course of this test, the BESS experienced six calibration events. We can see that, over time, both CALVS and CANVS, outperform the CCA model. The first few calibration events serve as training steps. Note that, CANVS performs better than CALVS because it is more accurate, and better positioned to fine tune the model, especially around the calibration events.

The corrective online learning capability of the above proposed approaches at calibration events is clearly an advantage, considering that the common practice in estimating a pack’s maximum capacity requires dedicated charge and discharge test cycles for the purpose of modeling. Of course, we can still benefit from such dedicated tests in the initial commissioning stage. But we no longer need to interrupt the normal operation of the BESS because the proposed approach can continuously update the models whenever a calibration event is encountered.

C. Optimal Energy Storage Operation

In this section, we consider the application of BESS in peak load shaving and compare the performance of the proposed CANVS model with the baseline CCA model using the PHIL testbed. The energy and power ratings of the BESS are assumed to be 240 kWh and 80 kW, respectively. The results are shown in Fig. 12. The peak load occurs at hour 19.

First consider the performance of the CCA model. A similar issue such as the one in Event 1 in Fig. 1 occurs at hour 19 which causes the battery to stop operating at a crucial peak hour; resulting in a poor peak load shaving performance. When the CCA model is used, the peak load is shaved by only 7 kW. In contrast, the proposed CANVS model performs well and
reduces the peak load by 18 kW. The reason is that when it comes to the CCA method, there is a big discrepancy between what the solution of the optimization problem assumes we will achieve versus what we actually achieve; and the discrepancy is due to poor model accuracy. In contrast, the CANVS model results in a less optimistic but far more realistic performance; thus, leading to a much better overall performance in practice.

D. Trade-offs in Using CALVS and CANVS

Recall that a critical parameter in the CANVS method is the number of LUT pairs. Increasing this number is expected to improve the model accuracy but at the expense of increasing computation time. This trade-off is studied in Fig. 13.

Fig. 13 (a) shows the results of the model accuracy for the CALVS model versus the CANVS model with different number of LUT pairs. The accuracy is quantified based on the final error at the end of a simple charge/discharge cycle. We can see that the CANVS approach achieves to a better accuracy than CALVS, when four or more pairs of \((C, V_{oc})\) are used in the LUT. However, the improvement in accuracy exhibits diminishing returns with more LUT pairs.

Fig. 13 (b) shows the computation time. Adding each LUT pair will increase computation time of the optimization, with an exponential relationship. Of course, this depends heavily on the solver, computer specifications, initial conditions, and optimality tolerance. Naturally, we have a trade-off between accuracy and computation time of the CANVS approach;

the binary expansion variable can be adjusted to enhance granularity, but at the cost of computation time.

E. Impact of Circuit Model Accuracy

Recall from Section II that the model in Fig. 4 was meant to capture the underlying device-level characteristics of the batteries while still maintaining relatively low complexity in the resulting BESS scheduling optimization problem. That of course raises the question on whether such model is indeed sufficiently accurate for the purpose of this study. To directly address this issue, in this section, we examine the accuracy of the battery model in the context of the peak-shaving problem.

In this regard, we use the battery hardware simulation tool in [52]. It is a detailed dynamic battery cell model implemented in Simulink [53]. Capacity, resistance, and capacitance are modeled dynamically. The model includes a physical model of the cell, heating from the internal resistance, and heat transfer characteristics when it is implemented as part of a pack.

The BESS operation in this case study is simulated for a duration of one month, i.e., 30 days. The operation schedule is updated on a daily basis. At midnight, the operation of the BESS is scheduled for the next 24 hours, once using the baseline CCA method and once using the proposed CALVS method. The resulting BESS operation schedule is then applied to the simulated hardware battery in [52]. The battery cells are selected so as to have 5% imbalance in their SoC and 5% variation in their internal resistance and capacitance to simulate battery imbalance and cell-to-cell variation. This matches what is commonly seen in practice, e.g., see [54].

The results for the case of the CCA method are presented in Fig. 14. We can see in Fig. 14(a) that the mismatch between the scheduled net load and the actual net load is considerable during the peak hours. As a result, peak load is shaved only at a fraction of what was scheduled. For example, peak load is shaved only at 80 kW on Day 9, even though it was scheduled to be shaved by 126 kW on this day. Similar results can be seen on several other days, as shown in Fig. 14(b), where the red bars are significantly less than the blue bars on several days. In summary, ignoring the underlying battery cell models results in major mismatch between the peak-load shaving that was intended and the peak-load shaving that was actually achieved.

Now, the main question is: how far can the proposed CALVS method resolve the above issue, even though the underlying battery cell model is relatively simple? Interestingly, the results are very promising, as shown in Fig. 15. In fact, we can see in Fig. 15(a) that the mismatch between the scheduled net load and the actual net load during the peak hours is much less for the CALVS method compared to the CCA method. For example, on Day 9, the peak load is shaved by 96 kW, which is exactly what was scheduled to be shaved. Similar patterns can be seen on almost all other days, as shown in Fig. 15(b), where the red bars are almost always at the same level as the blue bar. Although, there do exist few occasions, on Day 18 and Day 25, where the peak-load shaving that was achieved was less than the peak-load shaving that was intended. Therefore, we can conclude that, bringing a reasonably accurate cell-level battery model into the context of optimization problem formulation of
grid-tied BESS can significantly improve performance, while maintaining relatively low complexity for the system operator, if we use the methodologies that are developed in this paper.

Another way to compare the results in Figs. 14 and 15 is to look at the distribution of the shaved peak-load in each case. The results are shown in Figs. 16(a) and (b) for the CCA method and the CALVS method, respectively. We can see that, when the baseline CCA method is used, there is significant difference between the distribution of the peak-load shaving that was intended and the distribution of the peak-load shaving that was actually achieved. In contrast, when the CALVS method is used, there is only a slight difference between the distribution of the peak-load shaving that was intended and the distribution of the peak-load shaving that was actually achieved. Such slight difference is the cost that one may pay in order to maintain the analysis tractable so as to achieve an overall suitable optimization-based BESS scheduling method.

VI. CONCLUSIONS

We discussed the need for the optimization-based grid-tied BESS operation algorithms, such as peak-load shaving, to use per-cell battery models in order to provide reliable performance. A cell-level circuit model, consisting of a voltage source and internal resistance, was used to construct a pack model that is usable in a grid-integrated BESS optimization framework. Two methods were presented, CALVS and CANVS, to tackle the non-linear relationship between $V_{oc}$ and $C$. The trade-offs between the two methods, as well as their comparison with a traditional Coulomb counting approach, were explored. The models were trained using PHIL test data from a Lithium battery pack, and parameter estimation challenges were identified via a non-intrusive approach. The model was then validated in a peak load shaving experiment. It was observed that advanced modeling results in a far more realistic model and better performance for the entire pack; improving the operation of BESS in smart grid applications.

REFERENCES


APPENDIX A

BMS BASED CALIBRATED CCA METHOD ALGORITHM

When a BMS uses the CCA method, it also adds calibration points when the battery is completely full or completely empty. These points are highly non-linear and as such are not usable in an optimization framework. This algorithm is expressed below, where $C^{max}$ is the BMS estimate of total capacity.

**Algorithm 1 Calibrated CCCA**

```plaintext
\begin{align*}
t & \leftarrow \text{current timeslot} \\
SoC[t] & \leftarrow SoC[t - 1] + P[t - 1]/C^{max} \\
\text{if} \ V_o[t, n] < V_{min} \text{ then} \\
SoC[t] & \leftarrow 0\% \\
\text{else if} \ V_o[t, n] > V_{max} \text{ then} \\
SoC[t] & \leftarrow 100\% \\
\end{align*}
```

Zachariah Taylor (S’13-M’19) received the B.Sc. degree in electrical and computer engineering from California Baptist University, Riverside, CA, USA in 2013 and received the M.Sc. and Ph.D. degrees in electrical engineering from the University of California, Riverside, in 2014 and 2019, respectively. He has recently joined 174 Power Global in Irvine, CA, USA as a senior energy storage system engineer. His research interests include distributed resource optimization in power systems and operation, management, and modeling of energy storage systems.

Hossein Akhavan-Hejazi (S’12-M’17) received the M.Sc. degree in electrical engineering from the Amirkabir University of Technology, Tehran, Iran, in 2011 and the Ph.D. degree in electrical engineering from the University of California, Riverside, CA, USA, in 2016. He is currently an Assistant Research Faculty of Winston Chung Global Energy Center (WCGEC) and Assistant Adjunct Professor of Electrical and Engineering Department at Bourns College of Engineering UC Riverside. His research interests include optimization and stochastic analysis in electric power systems, power system operations and market analysis, operation management and modeling of energy storage systems, and data analytics in power systems applications.

Hamed Mohsenian-Rad (S’04-M’09-SM’14) received the Ph.D. degree in electrical and computer engineering from the University of British Columbia, Vancouver, BC, Canada, in 2008. He is currently an Associate Professor of electrical engineering with the University of California, Riverside, CA, USA. His research interests include modeling, data analysis, and optimization of power systems and smart grids. He was the recipient of the National Science Foundation CAREER Award, the Best Paper Award from the IEEE Power and Energy Society General Meeting, and the Best Paper Award from the IEEE Int. Conference on Smart Grid Communications. He was an Editor of the IEEE TRANSACTIONS ON SMART GRID and the IEEE POWER ENGINEERING LETTERS.