

State-Of-Health Estimation Pipeline for Li-ion Battery Packs with Heterogeneous Cells

Preet Gill¹, Dong Zhang², Luis D. Couto³, Chitra Dangwal¹, Sebastien Benjamin⁴, Wente Zeng⁵, Scott Moura¹

Abstract—The internal condition of lithium-ion batteries, in particular State-of-Health (SoH), needs careful monitoring to ensure safe and efficient operation. In this paper, we propose a hybrid online SoH estimation pipeline for series-connected heterogeneous cells. Implementing a single cell parameter estimation scheme for a battery pack with hundreds to thousands of cells is computationally intractable. This challenge is solved in this work using feature-based adaptive polling of cells with “extreme” parameter values. Furthermore, the electrical parameters for the polled cells are estimated using online recursive least squares with forgetting factor. The key novelty lies in accounting for the uncertain state dependence of the parameters. We use sparse Gaussian process regression to obtain the parameter bounds as a function of both SOC and temperature. The pipeline is validated through a simulation study, using experimental data from Li-NMC cells.

I. INTRODUCTION

Battery technology plays a pivotal role in achieving global energy sustainability and reduction of green house gas emissions. Due to high energy density, low self-discharging rate and long cycle life, li-ion batteries span a wide array of applications including consumer electronics, electric vehicles and grid-level energy storage. In recent years, much emphasis has been placed on developing real-time control and estimation of a battery’s internal states. However, with increasing performance demands, safe and efficient operation of Li-ion batteries remains a challenge.

Battery packs can contain up to hundreds or thousands of cells connected in series and parallel, depending on their application, such as electric vehicles and stationary grid storage. As such, it is crucial to monitor the internal states, such as state-of health (SoH), to ensure safety, performance and prolonged life. This is usually done by an advanced battery management system (BMS) that deploys real-time estimation and control schemes. However, SoH estimation in packs is

challenging due to: (i) scalability and computational cost, (ii) limited sensing, and (iii) complicated battery dynamics. An important fact often ignored during battery modeling is the time-varying electrical parameters. In practice, internal parameters, e.g. resistances and capacitances, depend nonlinearly on the cell’s temperature and SoC.

SoH estimation in batteries has been typically accomplished through two different routes. The first one consists in carrying out specific experimental procedures. This is algorithmically simple but requires one to stop battery operation and/or use specialized lab equipment [1]. The second route relies on models. It can be divided into data-driven and physics-based approaches for battery modeling. Examples of the former type include neural networks [2] and support vector machines [3] used to estimate SoH parameters like capacity. The drawback of these approaches is that they require large data sets and do not necessarily represent the battery physics. The latter physics-based approach describes battery dynamics using first-principles electrochemical models or electrical circuit analogies. Electrochemical models have been used to estimate SoH via the Levenberg–Marquardt method [4], Kalman filters [5], and Lyapunov methods [6], for instance. Although electrochemical models are relatively descriptive, they are difficult to identify given their sheer number of parameters [7] and associated identifiability issues [8]. They also tend to be complex nonlinear differential-algebraic equation systems. In contrast, equivalent circuit models (ECMs) are abstractions of the battery dynamics that strike a good balance between simplicity and ease of interpretation. Parameter estimation in ECMs includes nonlinear curve fitting [9], genetic algorithms [10], least-squares method [11] and Kalman filters [12], to name a few.

Most of the aforementioned efforts for SoH estimation have been centered around single cells, whereas few studies estimate parameters for interconnected cells in battery packs. For instance, Kalman filters were employed for combined SOC and SoH estimation in the context of equivalent circuit models [13], [14], [15], whereas global optimization approaches and particle filters [16], [17] also demonstrated high estimation accuracy. However, most of these approaches (i) lump the entire pack as one single virtual cell thereby losing critical cell heterogeneity information, or (ii) enforce one estimator on each in-pack cell and hence impose tremendous amount of computational burden. The estimation of parameters for every single cell in a battery pack of thousands of cells using highly non-linear and coupled dynamics is

*This work is sponsored by Total S.A. and Saft S.A..

¹Preet Gill, Chitra Dangwal and Scott J. Moura are with Department of Civil and Environmental Engineering, University of California, Berkeley, California, 94720, USA. {preetgill177, chitra_dangwal, smoura}@berkeley.edu

³Dong Zhang is with School of Aerospace and Mechanical Engineering, University of Oklahoma, Norman, Oklahoma 73019, USA. dzhang@ou.edu

²Luis D. Couto is with Department of Control Engineering and System Analysis, Université Libre de Bruxelles, B-1050 Brussels, Belgium. lcoutome@ulb.ac.be. He would like to thank the Wiener-Anspach Foundation for its financial support.

³Sebastien Benjamin is with Saft S.A.. Sebastien.BENJAMIN@saftbatteries.com

⁴Wente Zeng is with Total S.A.. wente.zeng@total.com

computationally intractable. The proposed SoH framework is scalable, since it derives the upper and lower bounds of all electrical parameters for all cells in a pack. Moreover, the framework uses sparse Gaussian process regressions to estimate parameter values from selected operating points to all the operating points of the battery pack. Given the aforementioned literature, this paper contributes:

- A novel framework to estimate parameter value bounds as a function of SOC and temperature, using sparse Gaussian process regression.
- An analysis of the voltage-based features for intelligent polling of representative cells in a battery pack.

The remainder of this paper is organized as follows. The equivalent circuit model for a single cell is developed in Section II. Next a brief motivation for the work is presented in Section III. The SoH framework for heterogeneous battery packs is developed in Section IV with Section IV-A outlining the SoH pipeline algorithm. Finally, a numerical assessment of the SoH pipeline is carried out in Section V.

Definition. Throughout the paper, *Representative Cells* are defined as cells in a battery pack whose parameters form the upper and lower bounds and enclose all the unmeasured parameters of all the cells in the pack. Furthermore, *Operating Point* is defined as the state of the battery pack - SoC ($\xi(t)$) and Temperature ($\Gamma(t)$) at any given time t . For the purposes of proper battery operation, the operating range is fixed to be $\xi \in [20\%, 100\%]$ and $\Gamma \in [10^\circ\text{C}, 60^\circ\text{C}]$.

II. MODEL DEVELOPMENT

This section reviews an equivalent-circuit model (ECM) for a single cell, which is then electrically interconnected with other cells to form a series arrangement.

A. ECM for Single Cell

The ECM for a single cell indexed k , consisting of a resistor in series with an $R-C$ pair and an open-circuit voltage (OCV), is defined by

$$\dot{\xi}_k(t) = \frac{1}{Q_k} I_k(t), \quad (1)$$

$$\dot{V}_{c,k}(t) = \frac{1}{C_k(\xi_k, \Gamma_k)} I_k(t) - \frac{1}{R_{2,k}(\xi_k, \Gamma_k) C_k(\xi_k, \Gamma_k)} V_{c,k}(t), \quad (2)$$

$$V_k(t) = V_{OCV}(\xi_k, t) + V_{c,k}(t) + I_k(t) R_{1,k}(\xi_k, \Gamma_k), \quad (3)$$

Here $\xi_k(t)$ represents the SOC for the k -th cell, $V_{c,k}(t)$ denotes the voltage across the $R_{2,k}-C_k$ pair, $R_{1,k}$ is the ohmic resistance, and Γ_k is the cell temperature. The electrical model parameters, namely $R_{1,k}$, $R_{2,k}$ and C_k are a function of cell SoC and temperature. It should be noted that cell capacity Q_k (given in Amp-hours, Ah) is a constant. The output (3) for any cell k in a pack expresses the relationship between the terminal voltage and open circuit voltage (V_{ocv} (a nonlinear function of SoC), voltage across resistor $R_{1,k}$, and voltage across the $R_{2,k}-C_k$ pair. Throughout the paper, positive current is specified for charging and negative for discharging.

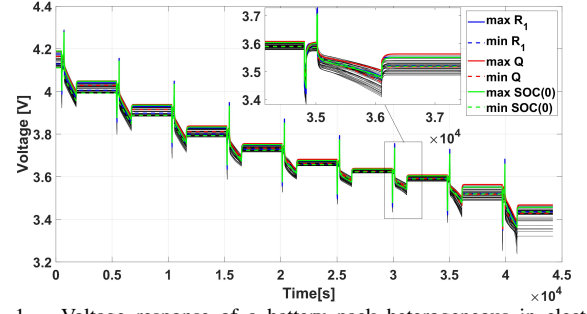


Fig. 1. Voltage response of a battery pack heterogeneous in electrical parameters R_1 , Q and initial SoC when given HPPC current profile.

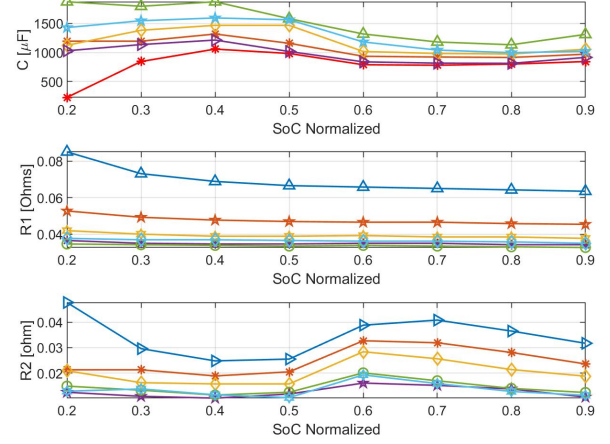


Fig. 2. Experimental data on Li-NMC cells. A reference performance test consisting of hybrid pulse power characterization was used to estimate true parameter values at temperatures $[10^\circ\text{C}, 60^\circ\text{C}]$ and $[20\%, 90\%]$ SOC. The plot shows R_1 , R_2 and C during discharge.

III. MOTIVATION

In this section, we illustrate cell heterogeneity with respect to electrical parameters R_1 , R_2 , C and Q using an open-loop simulation study [18], [19]. Without loss of generality, we consider five LiNiMnCoO₂/Graphite (Li-NMC) cells in series, each with a 2.8Ah nominal capacity. Through this simulation we will demonstrate that the “reductionist” approach prevalent in the literature of collapsing a battery pack with thousands of cells to two fixed representative cells – the “strongest” and “weakest” cells with respect to voltage, leads to incorrect conclusions.

In this formulation, all the cells have identical SoC-OCV relationship and the heterogeneity comes from:

- Difference in electrical parameters.
- Difference in SOC initialization.
- Difference in capacity Q .

For this demonstration, a modified Hybrid Pulse Power Characterization (HPPC) test cycle is applied at a temperature of 30°C to 50 cells in series as shown in Fig.1. The electrical parameters as a function of (ξ, Γ) are plotted in Fig 2.

For the demonstration, all three types of the aforementioned heterogeneity are applied to the pack of 50 cells, simultaneously, where the parameters and SoC are within $\pm 5\%$ of the fresh cell electrical parameters at 30°C , as shown in Fig. 2. As can be seen in Fig. 1, the voltage response of the battery pack V_{upper} and V_{lower} is not strictly

Algorithm 1: State-of-Health Pipeline

Inputs: Diagnostic current cycle**Outputs:** $R_1, R_2, C := f(\xi, \Gamma), Q$ **for** $\Gamma = 1 : T$ **do** **for** $\xi = 1 : N$ **do**

- *Run Diagnostic Cycle*
- *Poll representative cells using voltage features corresponding to each electrical parameter R_1, R_2, C , and Q*
- *Estimate $\underline{R}_1, \overline{R}_1, \underline{R}_2, \overline{R}_2, \underline{C}, \overline{C}, \underline{Q}, \overline{Q}$ using single cell recursive least square algorithm*

end**end**

- *Estimate parameters as $f(\xi, \Gamma) \forall \xi \in [20\%, 100\%]$ and $\Gamma \in [10^\circ\text{C}, 60^\circ\text{C}]$ using sparse Gaussian Process Regression*
-

correlated to the “extreme” cells - defined in literature as cells with min/max electrical parameters, and initial SoC. It can be observed that the cell with lowest ohmic resistance, and highest SoC- Q , doesn’t always have the highest voltage. Similarly, the cell with highest ohmic resistance and lowest SoC and Q doesn’t always have the lowest voltage response. Therefore, it can be concluded that using $V_{\min}$ and $V_{\max}$ as proxies for all the cells in a battery pack leads to incorrect bounds on parameter values. Moreover, in a large battery pack, the parameter values differ vastly. As such, collapsing the four dimensional parameter space to one dimension is incomplete. A new category of approaches is needed that takes into account (i) the heterogeneity of cells with respect to electrical parameters, and (ii) estimates parameters as a function of SoC and temperature.

IV. STATE-OF-HEALTH FRAMEWORK

In this section we introduce the State-of-Health pipeline that is divided into three parts: (i) feature analysis & cell polling, (ii) single cell electrical parameter estimation using recursive least squares (RLS), and (iii) sparse Gaussian process regression (sGPR) for estimating electrical parameters as a function of (ξ, Γ) where $\xi \in [20\%, 100\%]$ and $\Gamma \in [10^\circ\text{C}, 60^\circ\text{C}]$.

A. SoH Pipeline

This section provides a brief overview of the SoH pipeline as outlined in Algorithm 1. The objective is to estimate bounds on parameters R_1, R_2, C, Q across the pack, where the first three are functions of (ξ, Γ) . Broadly, the pipeline intelligently “polls” individual cells in the pack for characterization, and then uses sGPR to extrapolate parameter estimates across the pack.

The SoH pipeline uses output voltage measurement from a fixed charge-discharge pulse cycle that will be referred to as a “Diagnostic Cycle” for the remainder of the paper (see Appendix). The algorithm is divided into three main steps. These steps are executed for 9 fixed (ξ, Γ) operating

points. First, the voltage measurement from the diagnostic cycle is used to identify a set of representative cells using feature analysis discussed in Section IV-B. Second, using single cell online RLS with forgetting factor, the parameters of the representative cells are estimated. Finally, using sGPR, the upper bound and lower bound estimates of representative cells for 9 operating points are used to estimate an upper and lower bound parametric surface in $\mathbb{R}^3$, since the electrical parameters R_1, R_2 and C are functions of (ξ, Γ) . It should be noted that cell capacity is not a function of (ξ, Γ) .

B. Cell Polling and Feature Analysis

The cell to cell variation in a battery pack over its lifetime is often overlooked in approaches that estimate electrical parameters representative of the entire pack. However, as illustrated in Section III, the representative cells that correspond to the upper (max) and lower (min) bounds of the electrical parameters vary during operation and do not necessarily correlate with the best ($V_{\max}$) and worst ($V_{\min}$) performing cell. Moreover, the representative cells for each electrical parameter differ and seldom overlap as seen in Fig. 1. This motivates the need for an approach that can easily identify representative cells in real-time using the measured voltage signal. In [20], the authors developed “CiS” - Cells in Series approach to screen the cell to cell variation using voltage-based indexing as opposed to current capacity-based indexing. The CiS approach [20] was shown to consistently outperform existing capacity-based approaches.

In this section we discuss and demonstrate, via simulation, the voltage-based feature analysis used to intelligently poll representative cells corresponding to each electrical parameter. The slope of the charge/discharge voltage curves are known to provide insight on the electrical parameter values. This is widely used in industry for reference performance tests. The variation in resistance and capacity of the cells in a battery pack are directly related to variations in the voltage curve slope at the end of a pulse discharge and constant current discharge. As a result, the voltage response to a fixed pulse charge/discharge current cycle are leveraged to identify SoC-invariant voltage features that have high correlation coefficients with respect to (w.r.t.) electrical parameters. These correlation coefficients are then used to index the cells to identify the representative cells with maximum and minimum parameter values in the battery pack.

For the simulation study, a pack of 50 Li-NMC cells connected in series is used to generate voltage responses using the ECM model (1)-(3) with varying parameters, whose values are extracted from experimental data (Fig. 2). For this demonstration, the R_1, R_2, C and Q parameters are varied $\pm 5\%$ from the nominal value and the simulation is performed at a fixed temperature, 30°C . The goal is two fold: (i) identify voltage features that have the highest correlation factor w.r.t. to the parameters and (ii) identify the voltage features that are invariant w.r.t. SoC, i.e they have high correlation factors at all SoC levels ($\xi \in [10\%-100\%]$). The second goal is important because it provides the flexibility of intelligently

polling the representative cells at any given SoC during battery pack operation.

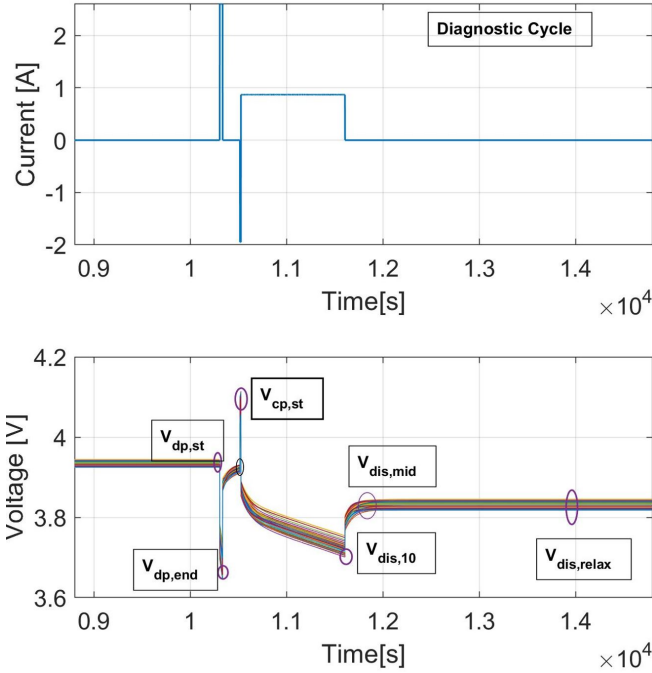


Fig. 3. Diagnostic pulse current profile and voltage response used throughout the paper for feature analysis, polling and the recursive least square algorithm. The bottom subplot indicates the voltage features used for intelligent polling.

Figure 3 shows the voltage response to the Diagnostic cycle – comprised of charge/discharge pulses – and the features used to perform the correlation analysis for all four electrical parameters. As can be seen in Fig. 4, the difference between discharge voltage at the start of the pulse and end of the pulse $\Delta V_{dp,st} = V_{dp,st} - V_{dp,end}$, and the voltage at the end of the pulse $V_{d,end}$ give correlation factors of -0.99 and -0.91 respectively, for ohmic resistance R_1 [Ω]. Other voltage features, namely the end of charge Voltage - $V_{cp,end}$ and end of relaxation voltage - $V_{dis,relax}$ give smaller correlation factors w.r.t. R_1 [Ω]. Furthermore, the features were analyzed at different SoC levels to check the dependence on SoC and $\Delta V_{dp,st}$ consistently had high correlation factor at all operating points. A similar two-phase feature analysis is conducted for the other electrical parameters $R_{2,k}$, C_k and Q_k , for $k \in [1, 50]$, to obtain voltage features that correspond to representative cells w.r.t. to all 4 electrical parameters. The features with correlation factor $\rho > 0.9$, corresponding to ECM parameters (R_2, C, Q) are: $(V_{dp,end} - V_{dp,st+1}, V_{dis,relax}, V_{dp,relax} - V_{dp,end+1})$ respectively. These identified features are used in Section V to intelligently poll the representative cells. Specifically, in Section IV-C we describe how polling involves estimating parameters of the representative cells using online Recursive Least Squares.

C. Recursive Least Squares (RLS) Model for Single Cell

In this section, we present the RLS algorithm for estimating a single cell's electrical parameters R_1 [Ω], R_2 [Ω], C [μF] and Q [Ah] for a fixed temperature and SoC. The RLS model

TABLE I
PARAMETER TRANSFORMATIONS

$(R_1, R_2, C) \rightarrow (\theta_1, \theta_2, \theta_3)$	$(\theta_1, \theta_2, \theta_3) \rightarrow (R_1, R_2, C)$
$\theta_1 = -\frac{1-2R_2C}{1+2R_2C}$	$R_1 = \frac{\theta_3 - \theta_2}{1 + \theta_1}$
$\theta_2 = -\frac{R_1 + R_2 + 2R_1R_2C}{1+2R_2C}$	$R_2 = \frac{2(\theta_3 + \theta_1\theta_2)}{\theta_1^2 - 1}$
$\theta_3 = -\frac{R_1 + R_2 - 2R_1R_2C}{1+2R_2C}$	$C = \frac{-1}{4} \frac{(\theta_1 + 1)^2}{\theta_3 + \theta_1\theta_2}$

is divided into two parts: (i) First we formulate the RLS model for estimating R_1 , R_2 and C along with V_{ocv} which is treated as unknown. (ii) We use the inverted SOC-OCV relationship for formulating another RLS model to estimate cell capacity Q [Ah]. The RLS model for estimating R_1 , R_2 and C is adopted from [21]. Equations (2) and (3) are linear in the parameters, since SoC and temperature are fixed.

The algorithm begins by formulating a parametric model, where we take the Laplace transform, apply the bilinear transform, and then use the inverse Z transform:

$$sV_c(s) = \frac{I_L(s)}{C} - \frac{1}{R_2C}V_c(s) \quad (4)$$

$$V_k(s) - V_{OCV}(s) = I_L(s) \left(R_1 + \frac{R_2}{1 + R_2Cs} \right) \quad (5)$$

$$s = \frac{2}{T} \cdot \frac{1 + z^{-1}}{1 - z^{-1}} \quad (6)$$

$$V_k(z^{-1}) - V_{OCV}(z^{-1}) = I_L(z^{-1}) \cdot P \quad (7)$$

where,

$$P = \frac{\left(\frac{R_1T + R_2T + 2R_1R_2C}{T + 2R_2C} + \frac{R_1T + R_2T - 2R_1R_2C}{T + 2R_2C} \right) z^{-1}}{\left(1 + \frac{T - 2R_2C}{T + 2R_2C} \right) z^{-1}} \quad (8)$$

Using the inverse Z transform, we obtain a linear in the parameters output equation given by

$$V_t(k) = \quad (9)$$

$$V_{ocv}(k) - \theta_1 V_{ocv}(k-1) + \theta_1 V_t(k-1) + \theta_2 I_L(k) + \theta_3 I_L(k-1) \approx (1 - \theta_1) V_{ocv}(k) + \theta_1 V_t(k-1) + \theta_2 I_L(k) + \theta_3 I_L(k-1) \quad (10)$$

A key assumption to further simplify (9) to (10) is $V_{ocv}(k-1) \approx V_{ocv}(k)$. Since we consider V_{ocv} as an unknown parameter, we assume that changes in V_{ocv} are negligible to facilitate RLS. This assumption does not significantly affect our results since our voltage features in Section IV-B are based on very short duration charge/discharge pulses that conserve total SoC. Moreover, this assumption allows us to estimate V_{ocv} , and not assume SoC-OCV to be constant over the lifetime of the battery. Table I below gives the forward and inverse transformations from $(R_1, R_2, C) \leftrightarrow (\theta_1, \theta_2, \theta_3)$. Finally, we have the RLS model of the form:

$$y(t) = \vartheta^T \phi(t) \quad (11)$$

where the regressor ϕ and parameters ϑ are given by

$$\phi = [1, V_t(k-1), I(k), I(k-1)] \quad (12)$$

$$\vartheta = [(1 - \theta_1)V_{oc}(k), \theta_1, \theta_2, \theta_3] \quad (13)$$

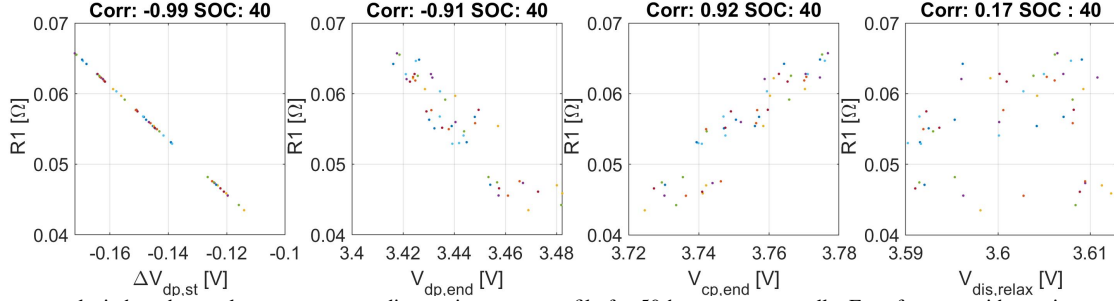


Fig. 4. Feature analysis based on voltage response to diagnostic current profile for 50 heterogeneous cells. Four features with varying correlation factors corresponding to R_1 at a fixed SOC are shown.

We use recursive least square optimization with forgetting factor to estimate $\hat{\vartheta}$ directly from (11). The parameters θ_1, θ_2 , and θ_3 are identified using,

$$y(k) = \phi(k) \vartheta(k) \quad (14)$$

$$K(k) = \frac{P(k-1)\phi(k)^T}{\phi(k)P(k-1)\phi(k)^T + \lambda} \quad (15)$$

$$\hat{\vartheta}(k) = \hat{\vartheta}(k-1) + K(k)[z(k) - \phi(k)\hat{\vartheta}(k-1)] \quad (16)$$

$$P(k) = \frac{I - K(k)\phi(k)^T P(k-1)}{\lambda} \quad (17)$$

where $\hat{\vartheta}(k)$ is the estimate of parameters ϑ_k in (13), $P(k)$ is the covariance matrix, $K(k)$ is the gain and $\lambda \in [0, 1]$ is the forgetting factor.

Finally, using a similar approach but in continuous time, we formulate an RLS model for estimating the cell capacity $Q[Ah]$ by first taking the Laplace of (1) to derive the transfer function from $I(t)$ to $\xi(t)$

$$s \cdot \xi(s) = \frac{-1}{Q} I(s) \quad (18)$$

We have a linear in the parameter system. However it involves a derivative of a measured signal. Therefore we apply a first order filter to both sides of (18):

$$\frac{1}{\Lambda(s)} = \frac{\lambda}{s + \lambda}, \quad (19)$$

which leads to

$$\underbrace{\left\{ \frac{s\lambda}{s+\lambda} \right\}}_{=\zeta(s)} Z(s) = \frac{-1}{Q} \underbrace{\left\{ \frac{\lambda}{s+\lambda} \right\}}_{=\rho(s)} I(s) \quad (20)$$

and we get a linear in the parameter form using the inverse Laplace transform given by (21) with estimation parameter $\theta = -1/Q$.

$$\xi(t) = \theta^T \rho(t) \quad (21)$$

Finally, we estimate Q using RLS with forgetting factor as defined in (14) - (17) where $\hat{\vartheta}_k$ is the estimate of θ which is $\frac{-1}{Q}$ in this step.

D. Sparse Gaussian Process Regression

In this section, we briefly introduce Gaussian Process Regression (GPR) and its variation – sparse Gaussian Process Regression (sGPR). Furthermore, we will formulate the

sparse GPR for estimation of electrical parameters as a function of (ξ, Γ) .

GPR is a powerful tool for non-parametric regression in high-dimensional spaces. One of the appealing features of GPR is that it provides uncertainty bounds around the estimates. A GPR is fully described by its mean and covariance, where the outputs of the modeled function are jointly distributed [22]. In essence, a GPR *a priori* describes the behavior of the function values using the following assumptions

$$y_i = f(x_i) + \varepsilon \quad (22)$$

where y_i and x_i are the scalar output and vector input $x_i \in \mathbb{R}^D$, and ε is drawn from $N(0, \sigma^2)$. Then the joint distribution is given by

$$p(y|X) = N(0, K(X, X) + \sigma^2 I) \quad (23)$$

where covariance matrix K is called the Kernel matrix. The joint posterior [23], given the Bayesian inference, can be defined as

$$p(f, f^*|y) = \frac{p(y|f)p(f, f^*)}{p(y)} \quad (24)$$

$$p(f^*|y) = \int \frac{p(y|f)p(f, f^*)}{p(y)} df \quad (25)$$

By marginalizing the latent variables, the joint prior distribution and independent likelihood probability are given by

$$p(f, f^*) = N \left(0, \begin{bmatrix} K_{f,f^*} & K_{f^*,f} \\ K_{f,f} & K_{f^*,f^*} \end{bmatrix} \right) \quad (26)$$

$$p(y|f) = N(f, \sigma^2 I) \quad (27)$$

Here, f and f^* are variables for the covariance calculation and I is the identity matrix. Finally, we can get a closed form for the GP predictive distribution by evaluating the integral in (25) to get

$$p(f^*|y) = N(\mu, \Sigma) \quad (28)$$

Here μ is the predicted output mean and Σ is the variance. In order to reduce the computational burden and large data demand of GPR, we adopted sparse GPR with the inducing variables approach from Titsias *et. al* [24]. Using this approach we can choose m inducing variables which are latent functions evaluated at some X_m inputs. These inputs X_m can then be identified as variational parameters by minimizing

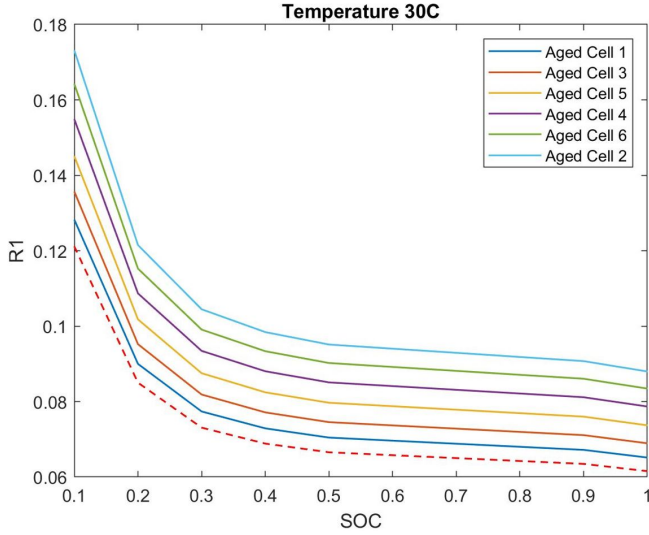


Fig. 5. Heterogeneity induced by uniformly perturbing the nominal R_1 parameter shown in dashed red line (from experimental data) of the ECM model as a $f(\xi)$ at 30°C for 6 cells in series (Ground truth values for validation). The process is repeated for all temperatures in operating range.

the distance between the exact posterior GP given by (28) and a variational approximation [24]. For this work, we use 3 inducing points (ξ) for a fixed Γ . More details of the approach are omitted here. Interested readers can refer to [24] for more details.

V. SIMULATION STUDY

In this section, the SoH pipeline is validated using numerical experiments on a pack of 6 Li-NMC cells connected in series, modeled using (1)-(3). The state-dependent electrical parameters used are obtained using experimental data on a beginning-of-life Li-NMC cell, as seen in Fig. 2. The steps outlined in Algorithm 1 are used to estimate bounds on the electrical parameters of representative cells, as a function of SoC and temperature given a fixed charge-discharge pulse diagnostic cycle applied to the battery pack.

First, we consider a battery pack of 6 cells connected in series where electrical parameters, R_1 , R_2 , C and Q , are uniformly perturbed between 5-7% from the fresh cell to induce heterogeneity in the battery pack. The output voltage measurement at 9 distinct operating points ($\{\xi \in \{30\%, 50\%, 80\% \} \times \{\Gamma \in \{20^\circ\text{C}, 30^\circ\text{C}, 40^\circ\text{C}\} \}$) is then used to intelligently poll the index of representative cells w.r.t. each electrical parameter as given in Algorithm 1, for the respective operating point, using the framework in Section IV-B. The electrical parameters of the representative cells are estimated using the RLS algorithm in Section IV-C. Finally, the estimated parameters are used to predict the parametric space as a function of SoC and temperature using sparse GPR as outlined in Section IV-D, for all the operating points.

For the purposes of demonstration, this section considers the parameter R_1 which is a function of (ξ, Γ) . Figure 5 shows the uniformly perturbed values of R_1 in a pack of 6 cells. It should be noted that Fig. 5 also illustrates the lack of correlation between spatial proximity of the cells in a

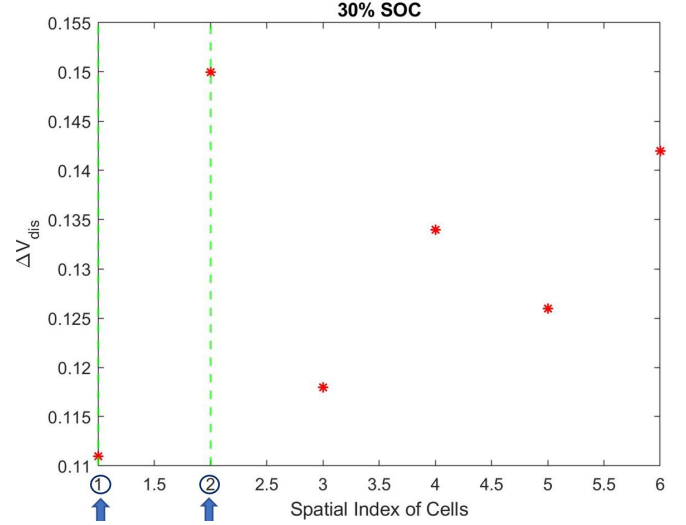


Fig. 6. Intelligent polling validation: Identified feature ΔV_{dis} used to poll representative cells from a heterogeneous pack of 6 cells in series. Here, the index of cells are on the x-axis and feature ΔV_{dis} is on y-axis.

pack and parameter values of the cells in the pack. As can be seen in Fig. 6, cell index 1 and 2 have the lowest and highest feature ($\Delta V_{dis} = V_{dp,st} - V_{dp,end}$) values, respectively, and hence are identified as representative cells for the entire pack for parameter R_1 . Since features from Section IV-B are invariant w.r.t. ξ , the representative cells w.r.t. R_1 for other operating points remains fixed. Figure 7 shows the estimation of R_1 for representative cell index 1 and 2 converging to the true value with high accuracy (RMSE = 0.002Ω). More importantly, this also verifies that the estimated upper bound and lower bound encloses all the unmeasured parameter values for all other cells in the pack. Finally, Fig. 8 shows the a cross-section of estimated parametric surface of $R_1 = f(\xi, \Gamma)$ for $\Gamma = 20^\circ\text{C}$, using sGPR. The solid blue curve represents the mean prediction of R_1 at $\Gamma = 20^\circ\text{C}$, whereas the shaded blue region shows the $\pm 95\%$ confidence interval around the estimates. This results confirms that the estimates of the parameter at selected operating points, in this case $\xi \in \{30^\circ\text{C}, 50^\circ\text{C}, 80^\circ\text{C}\}$, can be used to predict values at other operating points $\xi \in [20^\circ\text{C}, 90^\circ\text{C}]$, with uncertainty bounds around the estimates.

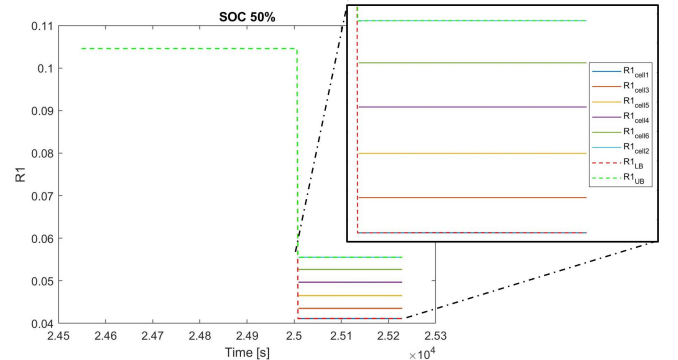


Fig. 7. Recursive least squares estimates of R_1 for the representative cells, indexed 1 and 2. The estimates of R_1 for these cells enclose all values of R_1 in the pack, thus verifying the polling approach using voltage feature correlation.

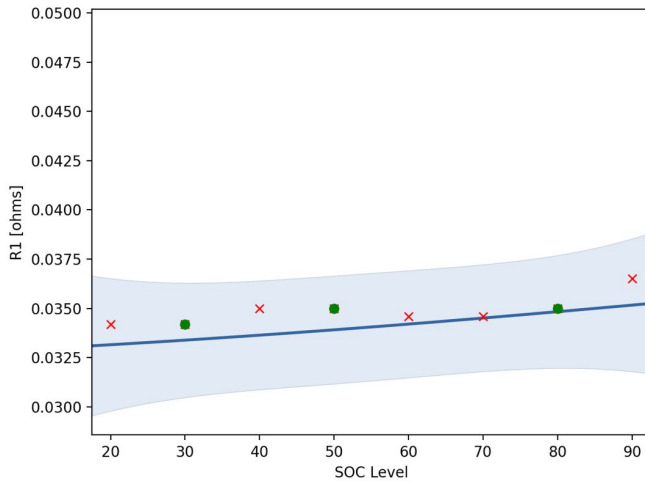


Fig. 8. sGPR results: Cross-Section of estimated $R_1 = f(\xi, \Gamma)$ at temperature $\Gamma = 20^\circ\text{C}$. Green points are sampled data and red crosses are true values of R_1 at un-sampled points.

VI. CONCLUSION

A State-of-Health (SoH) framework based on an equivalent-circuit model for heterogeneous cells in a lithium-ion battery pack is presented in this paper. The electrical parameters of the cells are considered as a function of SoC and temperature. A novel feature-based polling approach is developed to identify the index of bounding representative cells of the pack. Given a fixed operating point, i.e. SoC and temperature, the output of a charge-discharge pulse cycle is used to estimate electrical parameters via recursive least squares (RLS). The estimates from RLS are used to estimate the parameters as a function of SoC and temperature using sparse Gaussian Process Regression. The SoH framework is validated using numerical experiments based on electrical parameters obtained through experimental data acquired from a Li-NMC cell. An important feature of the proposed architecture is scalability, both in terms of the number of cells and operating points, since the number of representative cells are independent of the total number of cells in the battery pack and estimates at three operating points are used to estimate the entire parametric space. Furthermore, the efficacy of the pipeline design is validated numerically. Future work includes validation using experimental data for an aged battery pack of heterogeneous cells.

REFERENCES

- [1] Xinfan Lin, Hector E Perez, Shankar Mohan, Jason B Siegel, Anna G Stefanopoulou, Yi Ding, and Matthew P Castanier. A lumped-parameter electro-thermal model for cylindrical batteries. *Journal of Power Sources*, 257:1–11, 2014.
- [2] Duo Yang, Yujie Wang, Rui Pan, Ruiyang Chen, and Zonghai Chen. A neural network based state-of-health estimation of lithium-ion battery in electric vehicles. *Energy Procedia*, 105:2059–2064, 2017.
- [3] Duo Yang, Yujie Wang, Rui Pan, Ruiyang Chen, and Zonghai Chen. State-of-health estimation for the lithium-ion battery based on support vector regression. *Applied Energy*, 227:273–283, 2018.
- [4] Shriram Santhanagopalan, Qi Zhang, Karthikeyan Kumaresan, and Ralph E. White. Parameter estimation and life modeling of lithium-ion cells. *Journal of The Electrochemical Society*, 155(4):A345–A353, 2008.
- [5] S. K. Rahimian, S. Rayman, and R. E. White. State of charge and loss of active material estimation of a lithium ion cell under low earth orbit condition using kalman filtering approaches. *Journal of The Electrochemical Society*, 159(6):A860–A872, 2012.
- [6] Dong Zhang, Satadru Dey, Luis D Couto, and Scott J Moura. Battery adaptive observer for a single-particle model with intercalation-induced stress. *IEEE transactions on control systems technology*, 28(4):1363–1377, 2019.
- [7] S. Santhanagopalan, Q. Guo, P. Ramadass, and R. E. White. Review of models for predicting the cycling performance of lithium ion batteries. *Journal of Power Sources*, 156(2):620–628, 2006.
- [8] A. M. Bizeray, J. Kim, S. R. Duncan, and D. A. Howey. Identifiability and parameter estimation of the single particle lithium-ion battery model. *IEEE Transactions on Control Systems Technology*, 27(5):1862–1877, 2019.
- [9] Z. Pei, Zhen Peng Xiaoxu Zhao, Huimei Yuan, and Lifeng Wu. An equivalent circuit model for lithium battery of electric vehicle considering self-healing characteristic. *Journal of Control Science and Engineering*, 2018:5179758, 2018.
- [10] H. He, R.; Xiong, and J. Fan. Evaluation of lithium-ion battery equivalent circuit models for state of charge estimation by an experimental approach. *Energies*, 4:582–598, 2011.
- [11] D. Zhang, S. Dey, H. E. Perez, and S. J. Moura. Remaining useful life estimation of lithium-ion batteries based on thermal dynamics. In *2017 American Control Conference (ACC)*, pages 4042–4047, 2017.
- [12] Xiangwei Guo, Longyun Kang, Yuan Yao, Zhizhen Huang, and Wenbiao Li. Joint estimation of the electric vehicle power battery state of charge based on the least squares method and the kalman filter algorithm. *Energies*, 9(2):100, 2016.
- [13] Gregory L. Plett. Extended kalman filtering for battery management systems of lipb-based hev battery packs: Part 3. state and parameter estimation. *Journal of Power Sources*, 134(2):277–292, 2004.
- [14] Xu Zhang, Yujie Wang, Duo Yang, and Zonghai Chen. An on-line estimation of battery pack parameters and state-of-charge using dual filters based on pack model. *Energy*, 115:219–229, 2016.
- [15] Hongbin Ren, Yuzhuang Zhao, Sizhong Chen, and Taipeng Wang. Design and implementation of a battery management system with active charge balance based on the soc and soh online estimation. *Energy*, 166:908–917, 2019.
- [16] Xu Zhang, Yujie Wang, Chang Liu, and Zonghai Chen. A novel approach of battery pack state of health estimation using artificial intelligence optimization algorithm. *Journal of Power sources*, 376:191–199, 2018.
- [17] Jun Bi, Ting Zhang, Haiyang Yu, and Yanqiong Kang. State-of-health estimation of lithium-ion battery packs in electric vehicles based on genetic resampling particle filter. *Applied energy*, 182:558–568, 2016.
- [18] Dong Zhang, Luis D Couto, Preet S Gill, Sebastien Benjamin, Went Zeng, and Scott J Moura. Thermal-enhanced adaptive interval estimation in battery packs with heterogeneous cells. *IEEE Transactions on Control Systems Technology*, 2021.
- [19] Dong Zhang, Luis D Couto, Preet Gill, Sebastien Benjamin, Went Zeng, and Scott J Moura. Interval observer for soc estimation in parallel-connected lithium-ion batteries. In *2020 American Control Conference (ACC)*, pages 1149–1154. IEEE, 2020.
- [20] Leqiong Xie, Dongsheng Ren, Li Wang, Zonghai Chen, Guangyu Tian, Khalil Amine, and Xiangming He. A facile approach to high precision detection of cell-to-cell variation for li-ion batteries. *Energy - Scientific Reports*, 10:2045–2322, 2020.
- [21] Zhirun Li, Rui Xiong, Hao Mu, he Hongwen, and Chun Wang. A novel parameter and state-of-charge determining method of lithium-ion battery for electric vehicles. *Applied Energy*, 207, 06 2017.
- [22] Geoffrey E. Hinton and Carl E. Rasmussen. Evaluation of gaussian processes and other methods for non-linear regression. 1997.
- [23] Qi Wu, Rob Law, and Xin Xu. A sparse gaussian process regression model for tourism demand forecasting in hong kong. *Expert Syst. Appl.*, 39:4769–4774, 04 2012.
- [24] Michalis K. Titsias. Variational learning of inducing variables in sparse gaussian processes. In *AISTATS*, 2009.