Time-efficient identification of lithium-ion battery temperature-dependent OCV-SOC curve using multi-output Gaussian process

Kesen Fan[a], Yiming Wan[a,*], Zhuo Wang[a], Kai Jiang[b]

[a]*Key Laboratory of Image Processing and Intelligent Control, Ministry of Education, School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan, 430074, Hubei, China*
[b]*State Key Laboratory of Advanced Electromagnetic Engineering and Technology, School of Electrical and Electronic Engineering, Huazhong University of Science and Technology, Wuhan, 430074, Hubei, China*

Abstract

For lithium-ion batteries, the functional dependence of open circuit voltage (OCV) on state of charge (SOC) varies with temperature and aging, which plays a significant role in accurate SOC estimation and state of health monitoring. To identify the OCV-SOC curve at a given condition, OCVs usually need to be either measured by a time-consuming OCV test, or estimated with inevitable errors that eventually propagate into the identified OCV-SOC curve. In this paper, we investigate time-efficient identification of temperature-dependent OCV-SOC curve from current-voltage data, without measuring or estimating OCVs. In particular, we identify the complete OCV-SOC curve from data over a partial SOC range at a given temperature, by fusing available OCV-SOC curve data at other temperatures. In the proposed approach, a multi-output Gaussian process (MOGP) model is first built to capture correlations among OCV-SOC curves at different temperatures, and then used to construct the OCV-SOC curve at the given temperature. Using experimental datasets, our proposed approach reduces the root mean square error (RMSE) of OCV predictions by at least 29.4% compared to three existing methods. Besides, with the updated OCV-SOC

*Corresponding author

Email addresses: kesenfan@hust.edu.cn (Kesen Fan), ywan@hust.edu.cn (Yiming Wan), wangzhuo@hust.edu.cn (Zhuo Wang), kjiang@hust.edu.cn (Kai Jiang)

curve, the RMSE of SOC estimates is reduced by at least 14.0%, compared to using a non-updated OCV-SOC curve.

Keywords: Lithium-ion battery; equivalent circuit model; open circuit voltage; state of charge; multi-output Gaussian process.

1. Introduction

With recent developments of electrode materials for energy storage [1, 2, 3, 4, 5], the performance of lithium-ion batteries (LIBs) have been further improved in cycle life, charge/discharge efficiency, and energy density. Due to these advantages, LIBs become essential components in various applications such as portable devices, power grids, and hybrid or electric vehicles [6, 7]. To ensure safety and reliability, an advanced battery management system (BMS) is required. The OCV (open circuit voltage)-SOC (state of charge) relationship is vital to BMS due to its importance in SOC estimation [8, 9, 10, 11, 12] and state of health (SOH) monitoring [13, 14, 15, 16].

The OCV-SOC relationship has dependence on operating conditions, such as temperature and aging [17, 18]. Hence, it is necessary to update the OCV-SOC curve as the operating condition varies. For this purpose, the OCV data are usually directly measured by an incremental or low-current OCV test [19]. In most literature, the OCV-SOC curve is described by a polynomial, exponential, fractional calculus, support vector machine, or neural network model [10, 13, 20, 21]. Such an OCV-SOC curve model needs to be completely re-identified at a different operating condition. In particular, [22] proposes a polynomial OCV-SOC model whose dependence on aging is explicitly parameterized by just one parameter. By doing so, the model update for aging involves updating only one parameter in [22]. Instead of identifying an OCV-SOC curve at one given temperature or aging condition, another line of research aims at constructing a temperature/aging-dependent OCV-SOC model using the OCV data collected under a set of temperatures or aging conditions. In [23], a polynomial model is constructed to express OCV as a function of SOC and temperature. In [24], a single-output Gaussian process is learned from multiple OCV-SOC curves under different temperatures to describe the dependence of OCV on SOC and temperature. In [25], a two-dimensional lookup table is used to represent OCV as a function of SOC and SOH. With the OCV data covering only a partial SOC range, a Kalman filtering approach is proposed in [25] to update the above lookup

table by exploiting correlations between multiple OCV-SOC curves under different conditions.

All literature mentioned above rely on the availability of the OCV data. However, it often takes hours or even days for an OCV test to directly measure OCVs over a partial or the full SOC range. Only a few papers focus on avoiding such a time-consuming OCV test and investigate time-efficient identification of the OCV-SOC curve. For this purpose, a two-step scheme is proposed in [8, 9]: first estimate OCVs from current-voltage measurements over the full or a partial SOC range, and then identify a polynomial model for the corresponding full or partial OCV-SOC curve. However, inevitable errors of the estimated OCVs propagate into the identified OCV-SOC curve, which cannot be effectively addressed by the above two-step scheme in [8, 9].

Without measuring or estimating OCVs, [26] proposes a migrated equivalent circuit model (ECM) to identify the OCV-SOC curve directly from current-voltage data. Firstly, the SOC-dependent ECM for a fresh cell is established as a base model, including the functional dependence of OCV and ECM parameters on SOC. Then, the migrated ECM is built on the base model via linear transformation, and the introduced migration factors are determined using the current-voltage data at a different operating condition. The updated OCV-SOC curve is included in the identified migrated model. Although the above migration approach in [26] avoids the error propagation from estimated OCVs, its achievable performance is inherently limited by describing the effect of temperature and aging by linear transformation of the base model.

In this paper, we focus on time-efficient identification of the OCV-SOC curve using current-voltage data under a given temperature, without measuring or estimating OCVs. In particular, we identify the entire OCV-SOC curve over the full SOC range from data of a partial charging/discharging process at the given temperature, by fusing available OCV-SOC curve data at other temperatures. The proposed approach adopts a multi-output Gaussian process (MOGP) model to capture correlations among OCV-SOC curves at different temperatures, and predicts OCVs at the given temperature by the posterior means. The main advantages of our proposed approach are summarized below.

- It is more time-efficient, because it avoids using a time-consuming OCV test to collect OCV measurements as in [22, 24, 25].
- It achieves higher accuracy than existing methods that avoid OCV

tests in OCV-SOC curve identification. Different from the two-step scheme in [8, 9] that requires OCV estimates, our proposed approach does not need to estimate OCVs, thus gets rid of propagating OCV estimation errors into the identified OCV-SOC curve. Moreover, for accurately describing different degrees of correlations among OCV-SOC curves at various temperatures, the non-parametric MOGP model in our proposed approach is more flexible than the parametric migrated model in [26] with a fixed model structure.

The rest of this paper is organized as follows. Section 2 states the OCV-SOC curve identification problem. Section 3 briefly reviews the non-parametric MOGP learning. In Section 4, the proposed MOGP based OCV-SOC curve identification approach is presented. Identification results via experimental data are discussed in Section 5. The concluding remarks are given in Section 6.

Notations. $\mathrm{diag}\left(\begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix}\right)$ denotes a n-by-n diagonal matrix whose i-th diagonal element is b_i. For a matrix B, the element at the i-th row and j-th column of B is denoted as $[B]_{i,j}$. I_n denotes a n-by-n identity matrix, while $\mathbf{0}_n$ denotes a n-dimensional column vector.

2. Model description and problem statement

In this section, we first describe the first-order ECM, and then state the OCV-SOC curve identification problem to be solved.

2.1. The first-order ECM

Due to its simplicity and accuracy, the first-order ECM in Fig. 1 is widely adopted in battery applications [8]. The internal resistance R_s is used to represent the ohmic polarization phenomenon. The polarization resistance R_1 and polarization capacitance C_1 are used to describe electrochemical polarization and concentration polarization, which reflect the transient dynamics under current excitation. The OCV-SOC curve, R_s, R_1 and C_1 vary with temperature [27, 28].Based on the circuit theory, the first-order ECM at a fixed temperature is expressed as:

$$\frac{dV_1}{dt} = \frac{I}{C_1} - \frac{V_1}{R_1 C_1}, \tag{1a}$$

$$V_t = V_{oc}(z) - IR_s - V_1, \tag{1b}$$

where I is the load current with a positive value at discharge, V_t is the terminal voltage, V_1 denotes the overpotential voltage across R_1, $V_{oc}(z)$ represents OCV as a function of the SOC z.

Remark 1. *For the sake of notation simplicity, we do not consider the dependence of R_s, R_1, and C_1 on SOC in this paper, which suffices to give an accurate ECM in most cases. As mentioned in [29, 30, 31], these three ECM parameters generally vary with SOC only in the low SOC range (0-20%), but have no obvious dependence on SOC in the main operating range (20-100%). By following the method in [32], our proposed approach can be extended to the case of SOC-dependent ECM parameters without difficulty.*

To identify ECM parameters from sampled measurements, the continuous-time ECM in (1a) is transformed into a discrete-time model. Let T_s denote the sampling interval. By assuming constant current input over each sampling period $[(k-1)T_s, kT_s]$, it can be then derived from (1a) that

$$V_{1,k} = R_1 \left(1 - e^{-\frac{T_s}{\tau}}\right) I_{k-1} + e^{-\frac{T_s}{\tau}} V_{1,k-1}, \tag{2}$$

with $\tau = R_1 C_1$, $V_{1,k} = V_1(kT_s)$, $I_k = I(kT_s)$. By substituting (1b) into (2), we obtain the following regression model:

$$\begin{aligned} V_{t,k} &= \theta_{0,k} + \theta_1 V_{t,k-1} - \theta_2 I_{k-1} - \theta_3 I_k + \zeta_k, \\ &\approx \theta_{0,k} + \theta_1 V_{t,k-1} - \theta_2 I_{k-1} - \theta_3 I_k, \end{aligned} \tag{3}$$

where $V_{t,k} = V_t(kT_s)$ represents the sampled terminal voltage, $z_k = z(kT_s)$, $V_{oc,k} = V_{oc}(z_k)$, and the model coefficients $\{\theta_i\}_{i=0}^3$ are defined as

$$\theta_{0,k} = \theta_0(z_k) = (1 - \theta_1) V_{oc,k}, \tag{4a}$$

$$\theta_1 = e^{-\frac{T_s}{\tau}},\ \theta_3 = R_s, \tag{4b}$$

$$\theta_2 = R_1 (1 - \theta_1) - R_s \theta_1, \tag{4c}$$

$$\zeta_k = \theta_1 (V_{oc,k} - V_{oc,k-1}). \tag{4d}$$

Since $|z_k - z_{k-1}|$ is usually sufficiently small over a short sampling interval T_s, $|V_{oc,k} - V_{oc,k-1}|$ is close to zero [18]. Also considering that θ_1 is less than 1, ζ_k in (4d) is therefore sufficiently small to be negligible.

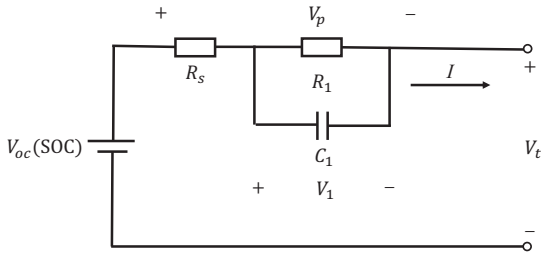

Figure 1: The first-order ECM

2.2. Problem statement

In this paper, we aim at identifying the complete OCV-SOC curve over the full SOC range from current-voltage data of a partial charging/discharging process at a given temperature T_1. In doing so, we avoid the time-consuming OCV test to measure OCVs, and get rid of estimating OCVs whose estimation errors propagate into the identified OCV-SOC curve.

In the above problem, the data over a partial SOC range at the given temperature T_1 are denoted by $\{\mathbf{I}^1, \mathbf{V}_t^1, \mathbf{z}^1\}$, where $\mathbf{I}^1 = \{I_k^1\}_{k=1}^{n_1}$ and $\mathbf{V}_t^1 = \{V_{t,k}^1\}_{k=1}^{n_1}$ are the current and voltage measurements, and $\mathbf{z}^1 = \{z_k^1\}_{k=1}^{n_1}$ represents the associated SOCs that are assumed known and can be obtained via an SOC estimation method. Assume that OCV tests at a limited number of temperatures $\{T_i\}_{i=2}^m$ other than T_1 are also performed offline. Hence the generated OCV-SOC curve data $\{\mathbf{z}^i, \mathbf{V}_{oc}^i\}_{i=2}^m$ over the full SOC range at temperatures $\{T_i\}_{i=2}^m$ are available, with $\mathbf{z}^i = \{z_k^i\}_{k=1}^{n_i}$ and $\mathbf{V}_{oc}^i = \{V_{oc,k}^i\}_{k=1}^{n_i}$ denoting the SOCs and the corresponding OCVs at temperature T_i, respectively. The above OCV-SOC curve identification problem will be solved by fusing the partial charging/discharging data $\{\mathbf{I}^1, \mathbf{V}_t^1, \mathbf{z}^1\}$ at temperature T_1 and the OCV-SOC curve data $\{\mathbf{z}^i, \mathbf{V}_{oc}^i\}_{i=2}^m$ at other temperatures $\{T_i\}_{i=2}^m$.

3. Preliminaries on multi-output Gaussian process

In this section, we briefly review MOGP, which is the basis to develop our proposed approach in the next section. More details of MOGP are referred to [33, 34].

MOGP is an extension of SOGP [35] by taking into account correlations among multiple outputs. Consider a vector-valued function $h(a) =$

$\begin{bmatrix} h^1(a) & h^2(a) & \dots & h^m(a) \end{bmatrix}^\top$ with m correlated outputs, where $h^i(a)$ is the i-th output with respect to input a. For the i-th output channel, we obtain the input data $\{a_l^i\}_{l=1}^{n_i}$ and the associated output data $\{h^i(a_l^i)\}_{l=1}^{n_i}$, with the data length n_i. For the sake of simplicity, we assumes that the input data of different output channels are identical with the same data length n, i.e.,

$$\{a_l^1\}_{l=1}^{n_1} = \{a_l^2\}_{l=1}^{n_2} = \cdots = \{a_l^m\}_{l=1}^{n_m} = \{a_l\}_{l=1}^{n}. \tag{5}$$

Note that this assumption can be relaxed, and MOGP can be easily generalized to the case where the input data of different channels are non-identical.

In MOGP, the m outputs of $h(a)$ are modeled as m dependent Gaussian processes with the covariance function

$$\operatorname{cov}(h^i(a_l), h^j(a_k)) = \rho_{ij}\kappa_a(a_l, a_k). \tag{6}$$

The above covariance function has a separable structure: ρ_{ij} captures the correlation between i-th and j-th output, and $\kappa_a(a_l, a_k)$ describes the covariance with respect to input a for a given output, which can be any valid kernel function with hyperparameters η. Other types of covariance functions for a general MOGP can be referred to [33]. Let $\boldsymbol{h}_i = \begin{bmatrix} h^i(a_1) & h^i(a_2) & \dots & h^i(a_n) \end{bmatrix}^\top$ denote the stacked i-th output data vector, with a_i defined in (5). Then $\boldsymbol{h} = \begin{bmatrix} \boldsymbol{h}_1^\top & \boldsymbol{h}_2^\top & \dots & \boldsymbol{h}_m^\top \end{bmatrix}^\top$ represents the output data vector of the vector-valued function $h(a)$.

With the covariance function (6), the joint probability distribution of the output data vector $\boldsymbol{h}$ is expressed as

$$p(\boldsymbol{h}) = \mathcal{N}(\mathbf{0}_{mn}, K_{\boldsymbol{h}}), \tag{7a}$$

$$K_{\boldsymbol{h}} = \boldsymbol{\rho} \otimes K_a + \Sigma \otimes I_n, \tag{7b}$$

where $\otimes$ represents the Kronecker product, $\Sigma = \operatorname{diag}\left(\begin{bmatrix} \sigma_1^2 & \sigma_2^2 & \dots & \sigma_m^2 \end{bmatrix}\right)$ with σ_i being the standard deviation of the measurement noise of the i-th output, K_a is a n-by-n covariance matrix whose (l,k)-th element is $[K_a]_{lk} = \kappa_a(a_l, a_k)$, and $\boldsymbol{\rho}$ is the label covariance matrix describing the correlations between different outputs with its (i,j)-th element being ρ_{ij}. Since $\boldsymbol{\rho}$ is a non-negative definite matrix, $\boldsymbol{\rho}$ can be expressed as follows using Cholesky factorization:

$$\boldsymbol{\rho} = LL^\top,\ L = \begin{bmatrix} L_{11} & & & \\ L_{21} & L_{22} & & \\ \vdots & \vdots & \ddots & \\ L_{m1} & L_{m2} & \dots & L_{mm} \end{bmatrix}, \tag{8}$$

with L being a lower-triangular matrix.

From (7), the joint distribution $p(\boldsymbol{h})$ is parameterized by $\{L_{ij}\}$ in (8) for $\boldsymbol{\rho}$, the noise variances $\{\sigma_i^2\}$ in Σ, and the hyperparameters η associated with K_a. This set of hyperparameters L, η, Σ can be obtained via maximum likelihood estimation with respect to $p(\boldsymbol{h})$ in (7), which is equivalent to minimizing

$$g_{\boldsymbol{h}}(L, \eta, \Sigma) = \boldsymbol{h}^\top K_{\boldsymbol{h}}^{-1} \boldsymbol{h} + \log(K_{\boldsymbol{h}}).$$

After determining the hyperparameters L, η, Σ, the joint distribution of $\boldsymbol{h}$ and the i-th output $h^i(a_*)$ at a given input a_* is

$$p\left(\begin{bmatrix} \boldsymbol{h} \\ h^i(a_*) \end{bmatrix}\right) = \mathcal{N}\left(\begin{bmatrix} \mathbf{0}_{mn} \\ 0 \end{bmatrix}, \begin{bmatrix} K_{\boldsymbol{h}} & K^i_{\boldsymbol{h},*} \\ K^i_{*,\boldsymbol{h}} & K^i_{*,*} \end{bmatrix}\right), \tag{9}$$

where $K_{\boldsymbol{h}}$ is defined in (7b), and

$$\begin{aligned} (K^i_{\boldsymbol{h},*})^\top = K^i_{*,\boldsymbol{h}} &= \begin{bmatrix} \rho_{i1} & \rho_{i2} & \dots & \rho_{im} \end{bmatrix} \otimes k_{*,\boldsymbol{h}}, \\ k_{*,\boldsymbol{h}} &= \begin{bmatrix} \kappa_a(a_*, a_1) & \kappa_a(a_*, a_2) & \dots & \kappa_a(a_*, a_m) \end{bmatrix}, \\ K^i_{*,*} &= \rho_{ii} \kappa_a(a_*, a_*). \end{aligned}$$

According to the Bayesian formula, the posterior distribution of $h^i(a_*)$ given $\boldsymbol{h}$ can be derived from (9), which is still Gaussian with its posterior mean and variance as

$$\mathbb{E}[h^i(a_*)|\boldsymbol{h}] = K^i_{*,\boldsymbol{h}} K_{\boldsymbol{h}}^{-1} \boldsymbol{h}, \tag{11a}$$

$$\operatorname{cov}[h^i(a_*)|\boldsymbol{h}] = K^i_{*,*} - K^i_{*,\boldsymbol{h}} K_{\boldsymbol{h}}^{-1} (K^i_{*,\boldsymbol{h}})^\top. \tag{11b}$$

It can be seen that the posterior mean of $h^i(a_*)$ in (11a) is a linear combination of $\boldsymbol{h}$, which means that MOGP uses available output data to estimate the i-th output $h^i(a_*)$ by exploiting correlations between different output channels.

4. OCV-SOC curve identification using MOGP

In this section, we identify the complete OCV-SOC curve from current-voltage data over a partial SOC range at a given temperature T_1. As depicted in Fig. 2, our proposed approach relies on MOGP learning and prediction to exploit both the current-voltage data $\{\mathbf{I}^1, \mathbf{V}_t^1, \mathbf{z}^1\}$ over a partial SOC range at temperature T_1 and the OCV-SOC curve data $\{\mathbf{z}^i, \mathbf{V}_{oc}^i\}_{i=2}^m$ over the full SOC

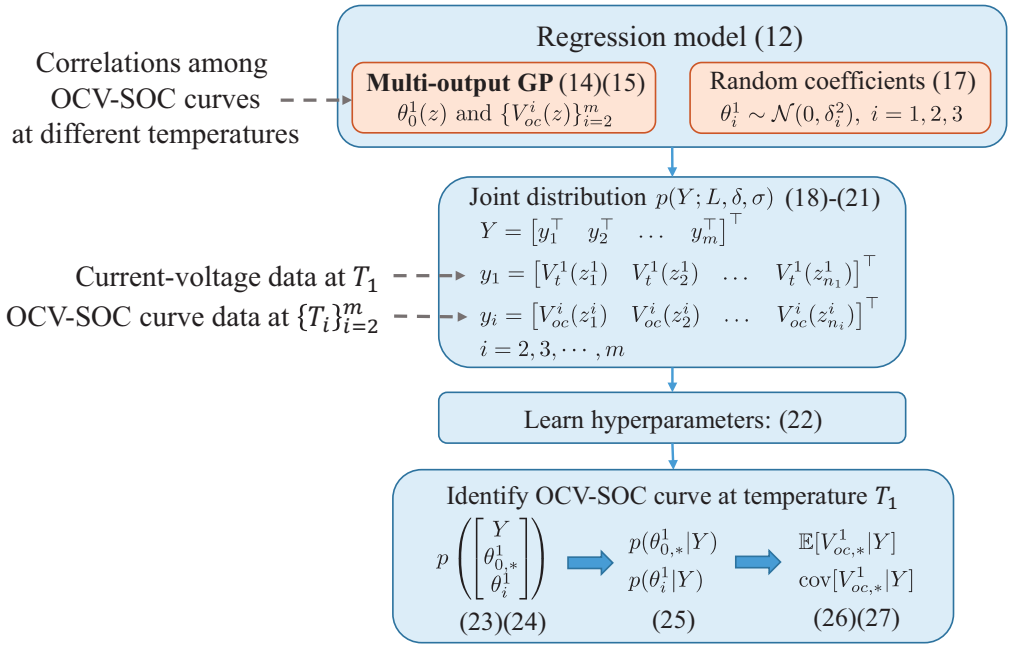

Figure 2: Flowchart of the proposed MOGP approach.

range at other temperatures $\{T_i\}_{i=2}^m$. It first identifies the model coefficients in (3), and then derives the OCV-SOC curve at temperature T_1. Our proposed approach achieves time-efficient identification of the OCV-SOC curve at temperature T_1 by avoiding the time-consuming OCV test.

4.1. Identification of model coefficients using MOGP

For the following derivations, the regression model (3) at temperature T_1 is rewritten as

$$V_{t,k}^1 = \theta_0^1(z_k) + \theta_1^1 x_{1,k}^1 + \theta_2^1 x_{2,k}^1 + \theta_3^1 x_{3,k}^1 + e_k^1, \tag{12}$$

with the white noise term $e_k^1 \sim \mathcal{N}(0, \sigma_1^2)$, and

$$x_{1,k}^1 = V_{t,k-1}^1,\ x_{2,k}^1 = -I_{k-1}^1,\ x_{3,k}^1 = -I_k^1. \tag{13}$$

The superscript "1" in (12) and (13) indicates variables at temperature T_1.

Since the OCV-SOC curves at different temperatures are similar in their shapes, the SOC-dependent parameter $\theta_0^1(z)$ at temperature T_1, defined as

$V_{oc}^1(z)$ multiplied with $1-\theta_1^1$ in (4a), is highly correlated with the OCV-SOC curves $\{V_{oc}^i(z)\}_{i=2}^m$ at other temperatures $\{T_i\}_{i=2}^m$. Therefore, we use an MOGP model to capture the correlations among $\theta_0^1(z)$ and $\{V_{oc}^i(z)\}_{i=2}^m$. These m outputs are described as zero-mean Gaussian processes, with their kernel function similar to (6), i.e.,

$$\kappa(z_l^i, z_k^j) = \rho_{ij} e^{-\frac{(z_l^i - z_k^j)^2}{2\delta_0^2}}. \tag{14}$$

Define $\theta_{0,k}^1 = \theta_0^1(z_k^1)$ and $V_{oc,k}^i = V_{oc}^i(z_k^i)$, with $k = 1, 2, \ldots, n_i$. From (14), we can obtain the covariance functions related to $\{V_{oc,j}^i\}$ and $\{\theta_{0,j}^1\}$:

$$\operatorname{cov}\left(V_{oc,l}^i, V_{oc,k}^j\right) = \kappa(z_l^i, z_k^j) + \mathbf{1}_{i,j}\mathbf{1}_{l,k}\sigma_i^2, \tag{15a}$$

$$\operatorname{cov}\left(\theta_{0,l}^1, V_{oc,k}^j\right) = \kappa(z_l^1, z_k^j),\ i, j = 2, \ldots, m, \tag{15b}$$

$$\operatorname{cov}\left(\theta_{0,l}^1, \theta_{0,k}^1\right) = \kappa(z_l^1, z_k^1), \tag{15c}$$

where $\mathbf{1}_{i,j}$ is the indicator function that is 1 only if i equals to j, and 0 otherwise, and σ_i^2 denotes the noise variance of the i-th output. The determination of hyperparameters in (14) and (15) will be discussed later. According to probability theory [36], the covariance matrix $\boldsymbol{\rho}$ defined in (8) determines the correlation coefficient matrix R whose (i,j)-th element

$$R_{ij} = \frac{\rho_{ij}}{\sqrt{\rho_{ii}\rho_{jj}}} \tag{16}$$

indicates the correlation between the OCV-SOC curves at temperatures T_i and T_j.

Besides the SOC-dependent parameter $\theta_0^1(z)$, the remaining model coefficients θ_1^1, θ_2^1 and θ_3^1 in (3) are described as three independent Gaussian random variables:

$$\theta_i^1 \sim \mathcal{N}(0, \delta_i^2),\ i = 1, 2, 3, \tag{17}$$

which are all assumed to be independent from $\theta_0^1(z)$ and the OCV-SOC curves $\{V_{oc}^i(z)\}_{i=2}^m$. Since $V_{t,k}^1$ in (12) is a linear combination of $\theta_{0,k}^1$ and $\{\theta_i^1\}_{i=1}^3$, we can derive the covariance function between $\{V_{oc,k}^i\}$ and $\{V_{t,k}^1\}$ based on (15):

$$\operatorname{cov}\left(V_{t,l}^1, V_{oc,k}^j\right) = \operatorname{cov}\left(\theta_{0,l}^1, V_{oc,k}^j\right) = \kappa(z_l^1, z_k^j), j = 2, 3, \ldots, m, \tag{18a}$$

$$\operatorname{cov}\left(V_{t,l}^1, V_{t,k}^1\right) = \kappa(z_l^1, z_k^1) + \sum_{i=1}^{3} \delta_i^2 x_{i,l}^1 x_{i,k}^1 + \mathbf{1}_{l,k}\sigma_1^2, \tag{18b}$$

where σ_1^2 is the variance of the measurement noise of V_t^1.

Let $y_1 = \begin{bmatrix} V_{t,1}^1 & V_{t,2}^1 & \dots & V_{t,n_1}^1 \end{bmatrix}^\top$ denote the measured terminal voltages at temperature T_1, and $y_i = \begin{bmatrix} V_{oc,1}^i & V_{oc,2}^i & \dots & V_{oc,n_i}^i \end{bmatrix}^\top$ the OCVs at temperature T_i, $i = 2, \cdots, m$. With (15a) and (18), we can obtain the joint distribution of the data vector $Y = \begin{bmatrix} y_1^\top & y_2^\top & \dots & y_m^\top \end{bmatrix}^\top$ as

$$p(Y) = \mathcal{N}\left(\mathbf{0}_N, K_{Y,Y}\right) \tag{19}$$

where $N = \sum_{i=1}^m n_i$ is the dimension of Y, and

$$K_{Y,Y} = \begin{bmatrix} K_{11} & K_{12} & \dots & K_{1m} \\ K_{21} & K_{22} & \dots & K_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ K_{m1} & K_{m2} & \dots & K_{mm} \end{bmatrix}, \tag{20}$$

K_{ij} is a n_i-by-n_j covariance matrix whose (l,k)-th element is

$$[K_{ij}]_{l,k} = \begin{cases} \operatorname{cov}\left(V_{t,l}^1, V_{t,k}^1\right), & i = j = 1, \\ \operatorname{cov}\left(V_{t,l}^1, V_{oc,k}^j\right), & i = 1, j = 2, \dots, m, \\ \operatorname{cov}\left(V_{oc,l}^i, V_{t,k}^1\right), & i = 2, \dots, m, j = 1, \\ \operatorname{cov}\left(V_{oc,l}^i, V_{oc,k}^j\right), & i, j = 2, \dots, m. \end{cases} \tag{21}$$

Note that covariances in the above equation are given in (15a) and (18).

With L defined in (8), $\delta = \begin{bmatrix} \delta_0, \delta_1, \delta_2, \delta_3 \end{bmatrix}$, and $\sigma = \begin{bmatrix} \sigma_1, \sigma_2, \dots, \sigma_m \end{bmatrix}$, hyperparameters L, δ, and σ can be obtained by maximum likelihood estimation. Maximizing the likelihood in (19) can be transformed into minimizing

$$g(L, \delta, \sigma) = Y^\top K_{Y,Y}^{-1} Y + \log\det K_{Y,Y}, \tag{22}$$

where the second term on the right-hand side represents the log-determinant of a matrix.

After determining the hyperparameters, given any input z_*^1, we can obtain the joint distribution of Y, $\theta_{0,*}^1 = \theta_0^1(z_*^1)$, and θ_i^1 as

$$p\left(\begin{bmatrix} Y \\ \theta_{0,*}^1 \\ \theta_i^1 \end{bmatrix}\right) = \mathcal{N}\left(\begin{bmatrix} \mathbf{0}_N \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} K_{Y,Y} & K_{Y,\theta_0} & K_{Y,\theta_i} \\ K_{\theta_0,Y} & K_{\theta_0,\theta_0} & 0 \\ K_{\theta_i,Y} & 0 & K_{\theta_i,\theta_i} \end{bmatrix}\right) \tag{23}$$

where

$$K_{\theta_0,\theta_0} = \kappa(z_*^1, z_*^1),\ K_{\theta_i,\theta_i} = \delta_i^2, \tag{24a}$$

$$K_{Y,\theta_i}^\top = K_{\theta_i,Y} = \begin{bmatrix} K_{\theta_i,y_1} & \mathbf{0}_{n_2}^\top & \dots & \mathbf{0}_{n_m}^\top \end{bmatrix}, \tag{24b}$$

$$K_{\theta_i,y_1} = \delta_i^2 \begin{bmatrix} x_{i,1}^1 & x_{i,1}^1 & \dots & x_{i,n_1}^1 \end{bmatrix}, \tag{24c}$$

$$K_{Y,\theta_0}^\top = K_{\theta_0,Y} = \begin{bmatrix} K_{\theta_0,y_1} & K_{\theta_0,y_2} & \dots & K_{\theta_0,y_m} \end{bmatrix}, \tag{24d}$$

$$K_{\theta_0,y_j} = \begin{bmatrix} \kappa(z_*^1, z_1^j) & \kappa(z_*^1, z_2^j) & \dots & \kappa(z_*^1, z_{n_j}^j) \end{bmatrix}, \tag{24e}$$

with $i = 1, 2, 3,\ j = 1, 2, \dots, m$. According to the Bayesian formula, we compute the posterior distributions of $\theta_{0,*}^1$ and $\{\theta_i^1\}_{i=1}^3$, which are all Gaussian with posterior means and variances as

$$\mathbb{E}[\theta_{0,*}^1|Y] = K_{\theta_0,Y} K_{Y,Y}^{-1} Y, \tag{25a}$$

$$\mathrm{cov}[\theta_{0,*}^1|Y] = K_{\theta_0,\theta_0} - K_{\theta_0,Y} K_{Y,Y}^{-1} K_{\theta_0,Y}^\top, \tag{25b}$$

$$\mathbb{E}[\theta_i^1|Y] = K_{\theta_i,Y} K_{Y,Y}^{-1} Y,\ i = 1, 2, 3, \tag{25c}$$

$$\mathrm{cov}[\theta_i^1|Y] = K_{\theta_i,\theta_i} - K_{\theta_i,Y} K_{Y,Y}^{-1} K_{\theta_i,Y}^\top. \tag{25d}$$

4.2. Identification of OCV-SOC curve from model coefficients

Next, we derive the OCV-SOC curve using the posterior means and variances in (25). According to (4a), $V_{oc,*}^1 = V_{oc}^1(z_*^1)$ nonlinearly depends on $\theta_{0,*}^1 = \theta_0^1(z_*^1)$ and θ_1^1, i.e.,

$$V_{oc,*}^1 = \frac{\theta_{0,*}^1}{1 - \theta_1^1}. \tag{26}$$

By performing the first-order Taylor expansion on (26), we can obtain the approximate posterior mean and variance of $V_{oc,*}^1$ as

$$\mathbb{E}[V_{oc,*}^1|Y] \approx \frac{\mu_{0,*}^1}{1 - \mu_1^1}, \tag{27a}$$

$$\mathrm{cov}[V_{oc,*}^1|Y] \approx \frac{\mathrm{cov}[\theta_{0,*}^1|Y]}{(1 - \mu_1^1)^2} + \frac{(\mu_{0,*}^1)^2 \mathrm{cov}[\theta_1^1|Y]}{(1 - \mu_1^1)^4}, \tag{27b}$$

with $\mu_{0,*}^1 = \mathbb{E}[\theta_{0,*}^1|Y]$ and $\mu_1^1 = \mathbb{E}[\theta_1^1|Y]$.

As can be seen from (25a), (25c), and (27a), the posterior mean of $V_{oc,*}^1$ has nonlinear dependence on $Y = \begin{bmatrix} y_1^\top & y_2^\top & \dots & y_m^\top \end{bmatrix}^\top$. This means that the OCV-SOC curve at temperature T_1 is identified by fusing the measured terminal voltages $y_1 = \begin{bmatrix} V_{t,1}^1 & V_{t,2}^1 & \dots & V_{t,n_1}^1 \end{bmatrix}^\top$ at temperature T_1 and the OCVs $\{y_i = \begin{bmatrix} V_{oc,1}^i & V_{oc,2}^i & \dots & V_{oc,n_i}^i \end{bmatrix}^\top\}_{i=2}^m$ at other temperatures $\{T_i\}_{i=2}^m$.

4.3. Summary of the proposed MOGP approach

The proposed MOGP based OCV-SOC curve identification algorithm is summarized below.

Inputs:

- The partial charging/discharging data $\{\mathbf{I}^1, \mathbf{V}_t^1, \mathbf{z}^1\}$ at temperature T_1.
- The OCV-SOC curve data $\{\mathbf{z}^i, \mathbf{V}_{oc}^i\}_{i=2}^m$ over the full SOC range at temperatures $\{T_i\}_{i=2}^m$.

Step 1. Learn hyperparameters

1a) Describe the model coefficient $\theta_0^1(z)$ and the OCV-SOC curves $\{V_{oc}^i(z)\}_{i=2}^m$ as an MOGP with zero means and covariance function (14); and regard $\{\theta_i^1\}_{i=1}^3$ as three mutually independent Gaussian random variables as described in (17);

1b) Determine the hyperparameters L, δ, and σ by minimizing (22), where L is defined in (8), $\delta = \left[\delta_0, \delta_1, \delta_2, \delta_3\right]$ consists of δ_0 in (14) and $\{\delta_i\}_{i=1}^3$ in (17), $\sigma = \left[\sigma_1, \sigma_2, \ldots, \sigma_m\right]$ includes σ_i used in (15a).

Step 2. Calculate OCV-SOC relationship at temperature T_1

2a) Compute the posterior means and variance functions of $\theta_0^1(z_*)$ and $\{\theta_i^1\}_{i=1}^3$ according to (25) , respectively.

2b) Compute the approximated posterior mean and variance function of $V_{oc}^1(z_*)$ according to (27).

4.4. Comparisons and discussions

In this subsection, our proposed MOGP approach will be compared with relevant methods in [8, 9, 24].

The two-step scheme in [8, 9] first estimates OCVs from current-voltage data, and then identifies a polynomial OCV-SOC model. The inevitable errors of OCV estimates in the first step eventually propagate into the identified OCV-SOC curve in the second step. In contrast, our proposed approach does not need to estimate OCVs at the given temperature T_1, but directly identifies the OCV-SOC curve with a nonparametric MOGP model. This is achieved by fusing the partial charging/discharging current-voltage data at temperature T_1 and the OCV-SOC curve data at other temperatures.

The SOGP approach in [24] relies on direct measurements of OCVs, and describes the functional dependence of OCV on SOC and temperature as a

single-output Gaussian process with two inputs. The kernel function of the adopted SOGP model is

$$\operatorname{cov}\left(V_{oc,l}^{i}, V_{oc,k}^{j}\right) = \lambda e^{-\frac{(T_i - T_j)^2}{2\delta_T^2}} e^{-\frac{(z_l^i - z_k^j)^2}{2\delta_z^2}}. \tag{28}$$

This implies that two OCV-SOC curves at temperatures T_i and T_j have a higher correlation when the temperature difference $|T_i - T_j|$ becomes smaller. However, such an implication may not be true in practice. As will be illustrated in Section 5.2, the above implication does not hold in the adopted experimental datasets, hence the SOGP approach results in a poor identification performance. In contrast, our proposed MOGP approach does not need OCV measurements obtained via a time-consuming OCV test, but identifies the OCV-SOC curve from current-voltage data. In particular, unlike the kernel function (28) of the SOGP approach, the kernel function (14) of our MOGP approach relies on ρ_{ij} to describe the correlation of OCV-SOC curves at temperatures T_i and T_j. The value of ρ_{ij} is learned from data, hence it is not necessarily large when the temperature difference $|T_i - T_j|$ is small.

5. Identification results using experimental datasets

This section presents the identification results of the proposed MOGP approach using experimental data of A123 LiFePO4 lithium-ion cells released by Center for Advanced Life Cycle Engineering (CALCE) at University of Maryland [37, 38]. The datasets are generated by a low-current OCV test and three dynamic tests including Federal Urban Driving Schedule (FUDS), US06 Highway Driving Schedule (US06) and Dynamic Stress Test (DST) at multiple fixed temperatures. The OCV-SOC curves obtained from low-current OCV tests are treated as the ground truth. We use the ground-truth OCV-SOC curve data over the full SOC range at $T_2 = 40°\text{C}$ and $T_3 = -10°\text{C}$ together with the FUDS current-voltage data over the 80-93% SOC range at $T_1 = 0°\text{C}$ to identify the entire OCV-SOC curve at $T_1 = 0°\text{C}$. The FUDS current-voltage data at temperature $T_1 = 0°\text{C}$ have $n_1 = 125$ samples, while the number of samples of the ground-truth OCV-SOC curve data at temperatures $T_2 = 40°\text{C}$ and $T_3 = -10°\text{C}$ are $n_2 = 1050$ and $n_3 = 127$, respectively.

Note that it takes 1050 seconds to collect data over the 80-93% SOC range in the FUDS test for our proposed approach. In comparison, to directly

measure OCVs over the full SOC range, the low-current OCV test in [19] takes 15276 seconds, which is significantly longer.

To illustrate the advantages mentioned in Section 4.4, our proposed MOGP approach will be compared with the two-step scheme in [8, 9] and the SOGP approach in [24]; then, we will show that the identified OCV-SOC curve using our proposed MOGP approach results in more accurate SOC estimation.

5.1. Comparison with the two-step scheme

In the above setting, only partial charging/discharging current-voltage data are provided at temperature T_1. As such, the two-step scheme in [8, 9] cannot be directly applied because it requires data covering the full SOC range. For a fair comparison with our proposed approach, the two-step scheme is modified as follows:

1) Estimating OCVs over a partial SOC range at temperature T_1: with $\theta_0^1(z)$ polynomially parameterized as

$$\theta_0^1(z_k^1) = c_d(z_k^1)^d + \cdots + c_1 z_k^1 + c_0, \tag{29}$$

the model coefficients $c_d, \cdots, c_1, c_0$ and $\{\theta_i^1\}_{i=1}^3$ in (12) can be identified using data $\{\mathbf{I}^1, \mathbf{V}_t^1, \mathbf{z}^1\}$ over the given partial SOC range; then, the corresponding OCV estimates can be calculated as

$$\hat{V}_{oc,k}^1 = \frac{\hat{\theta}_0^1(z_k^1)}{1-\hat{\theta}_1^1} = \frac{\hat{c}_d(z_k^1)^d + \cdots + \hat{c}_1 z_k^1 + \hat{c}_0}{1-\hat{\theta}_1^1}$$

according to (4), with $\{\hat{c}_i\}_{i=0}^d$ and $\{\hat{\theta}_i^1\}_{i=1}^3$ being the estimated model coefficients.

2) Identifying the OCV-SOC curve at temperature T_1: an MOGP model is first built by fusing the estimated OCV-SOC curve data $\{z_k^1, \hat{V}_{oc,k}^1\}_{k=1}^{n_1}$ over a partial SOC range at temperature T_1 and the OCV-SOC curve data $\{\mathbf{z}^i, \mathbf{V}_{oc}^i\}_{i=2}^m$ over the full SOC range at other temperatures $\{T_i\}_{i=2}^m$, with $\mathbf{z}^i = \{z_k^i\}_{k=1}^{n_i}$ and $\mathbf{V}_{oc}^i = \{V_{oc,k}^i\}_{k=1}^{n_i}$ denoting the SOCs and corresponding OCVs at temperature T_i, respectively; then, the OCV-SOC curve over the full SOC range at temperature T_1 is identified by using the posterior means derived from the constructed MOGP model. Although the above MOGP learning and prediction is similar to our proposed approach summarized in Section 4.3, the parameters $\{\theta_i^1\}_{i=1}^3$ in (12) no longer need to be considered here, because the OCV estimates at temperature T_1 are used in this step.

In the following comparison, we implement

1) LS-1-MOGP: the two-step scheme described above, with $d = 1$ in (29);
2) LS-2-MOGP: the two-step scheme described above, with $d = 2$ in (29);
3) Our proposed MOGP approach summarized in Section 4.3.

The MOGP learning and prediction is implemented by using the `GPML` toolbox [39]. To cope with local minimum during the hyperparameter optimization, we use multiple random initial guesses, and then select the set of learned hyperparameters which gives the best performance. However, the lower triangular matrix L in (8) includes 6 hyperparameters, which requires a great number of initial guesses. To address this issue, the initial guess for L with only one tuning parameter $\alpha \in [0, 1]$ is designed, which consists of the following two steps. In the first step, a simple form of the initial guess for the label covariance matrix $\boldsymbol{\rho}$ in (8) is adopted:

$$\boldsymbol{\rho} = \alpha E_m + (1 - \alpha) I_m, \tag{30}$$

where E_m is a m-by-m matrix with all elements equal to 1 and I_m is a m-by-m identity. It should be noted that a larger $\alpha \in [0, 1]$ indicates stronger correlation among OCV-SOC curves at different temperatures. In the second step, the initial guess for L is obtained by applying Cholesky factorization in (8) for $\boldsymbol{\rho}$ in (30). In our proposed MOGP approach, the learned hyperparameters are

$$\begin{aligned} L &= \begin{bmatrix} 2.52 & 0 & 0 \\ 1.66 \times 10^{-2} & 1.51 \times 10^{-2} & 0 \\ 2.13 \times 10^{-3} & 2.95 \times 10^{-2} & 8.61 \times 10^{-2} \end{bmatrix}, \\ \delta &= \begin{bmatrix} 426.83 & 14.29 & 1.33 & 1.39 \end{bmatrix}, \\ \sigma &= \begin{bmatrix} 2.30 \times 10^{-3} & 2.30 \times 10^{-3} & 2.30 \times 10^{-3} \end{bmatrix}. \end{aligned}$$

For LS-1-MOGP, the estimated partial OCV-SOC curve over the 80-93% SOC range in the first step is $V_{oc}^1(z_k^1) = 0.079 z_k^1 + 3.239$, and the hyperparameters learned in the second step are

$$\begin{aligned} L &= \begin{bmatrix} 4.36 & 0 & 0 \\ 4.41 & 0.03 & 0 \\ 3.11 & -0.90 & 0.03 \end{bmatrix}, \ \delta_0 = 0.063, \\ \sigma &= \begin{bmatrix} 1.47 \times 10^{-3} & 1.47 \times 10^{-3} & 1.47 \times 10^{-3} \end{bmatrix}. \end{aligned}$$

For LS-2-MOGP, the estimated partial OCV-SOC curve in the first step is $V_{oc}^1(z_k^1) = -0.039(z_k^1)^2 + 0.129z_k^1 + 3.226$, and the hyperparameters learned in the second step are

$$L = \begin{bmatrix} 4.44 & 0 & 0 \\ 4.50 & 0.02 & 0 \\ 3.29 & -0.88 & 0.02 \end{bmatrix},\ \delta_0 = 0.063,$$
$$\sigma = \begin{bmatrix} 1.48 \times 10^{-3} & 1.48 \times 10^{-3} & 1.48 \times 10^{-3} \end{bmatrix}.$$

Table 1: RMSEs (unit: volt) of identified OCV-SOC curve using the different approaches

proposed MOGP	LS-1-MOGP	LS-2-MOGP	SOGP
0.0156	0.0221	0.0237	0.0496

The identified OCV-SOC curves of three different approaches are shown in Fig. 3, with their associated RMSEs listed in Table 1. It can be seen that our proposed MOGP approach reduces the RMSE by at least 29.4% compared to the two-step methods. This is because the two-step methods introduce inevitable errors in the OCV estimates in the first step, but do not explicitly address them in the second step. In contrast, our proposed MOGP approach avoids estimating OCVs, and identifies the complete OCV-SOC curve directly from current-voltage data.

5.2. Comparison with the SOGP approach

Our proposed MOGP approach is further compared with the SOGP approach in [24] to illustrate the advantage explained in the third paragraph of Section 4.4. To implement the SOGP approach in [24], we assume that the ground-truth OCV-SOC curve over the 80-93% SOC range at temperature $T_1 = 0°\text{C}$ is available without any estimation errors. After that, the SOGP approach regards V_{oc} as an SOGP whose inputs consist of SOC and temperature. The SOGP model is constructed with the kernel function in (28), using the partial OCV-SOC curve data over the 80-93% SOC range at temperature $T_1 = 0°\text{C}$ and the complete OCV-SOC curve data over the full SOC range at temperatures $T_2 = 40°\text{C}$ and $T_3 = -10°\text{C}$. The hyperparameters learned for the SOGP model are $\delta_z = 0.05,\ \delta_T = 335.43,\ \lambda = 6.49,\ \sigma = 1.4 \times 10^{-3}$.

The identified OCV-SOC curves of these two approaches are shown in Fig. 4, with their RMSEs listed in Table 1. It can be seen that the proposed

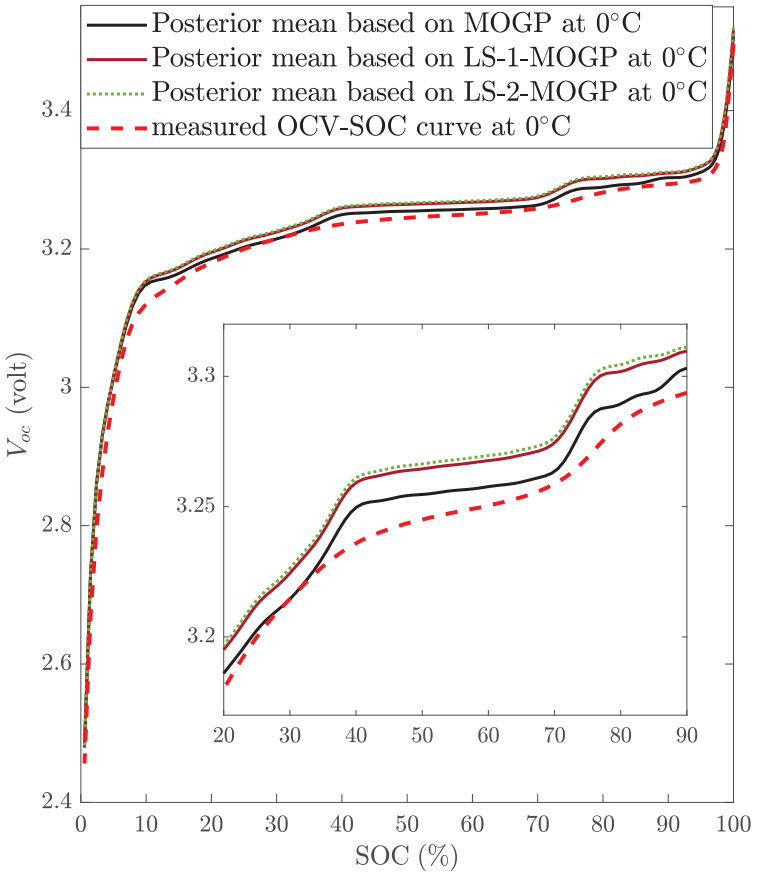

Figure 3: Identified OCV-SOC curves at temperature $T_1 = 0°\text{C}$ using the proposed MOGP approach and the two-step scheme (including LS-1-MOGP and LS-2-MOGP)

MOGP approach reduces the RMSE by 68.5% compared to the SOGP approach. The reason can be explained by the correlation coefficient matrix R in (16) learned from data. For the proposed MOGP approach,

$$R = \begin{bmatrix} 1 & 0.74 & 0.02 \\ 0.74 & 1 & 0.24 \\ 0.02 & 0.24 & 1 \end{bmatrix}$$

is obtained, and $R_{12} > R_{13}$ (R_{ij} represents the (i,j)-th element of R) implies that the OCV-SOC curves at $T_1 = 0°\text{C}$ and $T_2 = 40°\text{C}$ have a higher correlation than those at $T_1 = 0°\text{C}$ and $T_3 = -10°\text{C}$. This is consistent with the

ground-truth OCV-SOC curves in Fig. 4, i.e., the ground-truth OCV-SOC curves at $T_1 = 0°\text{C}$ and $T_2 = 40°\text{C}$ have a higher similarity than those at $T_1 = 0°\text{C}$ and $T_3 = -10°\text{C}$. However, such an observation is not illustrated by the correlation coefficient matrix

$$R = \begin{bmatrix} 1 & 0.999943 & 0.999996 \\ 0.999943 & 1 & 0.999911 \\ 0.999996 & 0.999911 & 1 \end{bmatrix}$$

learned in the SOGP approach. The essential reason for the above matrix R is due to the kernel function adopted by the SOGP approach, as explained in the third paragraph of Section 4.4.

5.3. Application of the proposed approach in SOC estimation

To illustrate the necessity of updating the temperature-dependent OCV-SOC curve, we identify ECMs at temperature $T_1 = 0°\text{C}$ with different assumptions on the OCV-SOC curve:

1) Baseline ECM-ID (ID is short for "identification"): the ECM parameters are identified via the popular least squares method [40] using the ground-truth OCV-SOC curve at $T_1 = 0°\text{C}$; this case uses the most accurate knowledge about the OCV-SOC curve, thus serves as the baseline;
2) ECM-ID with non-updated OCV-SOC curve: the ECM parameters are identified when using the ground-truth OCV-SOC curve at $T_3 = -10°\text{C}$ rather than $T_1 = 0°\text{C}$; this is the case that the OCV-SOC curve is not updated after the temperature varies from T_3 to T_1;
3) ECM-ID with updated OCV-SOC curve: the ECM parameters and the OCV-SOC curve are jointly identified by our proposed MOGP approach; this is the case that the OCV-SOC curve is updated after the temperature changes to T_1.

The obtained ECM parameters R_s, R_1 and C_1 of the above three models are shown in Table 2. These three identified ECMs are used to perform SOC estimation under the FUDS, US06, and DST conditions, all at temperature $T_1 = 0°\text{C}$. By following [8], we choose extended Kalman filtering (EKF) as the SOC estimation algorithm for the above three ECMs. We set the initial estimated state to be $\hat{z}_0 = 0.9$ and $\hat{V}_{1,0} = 0$. In the EKF algorithm, the covariance matrices of the initial estimated state, the process noise, and the

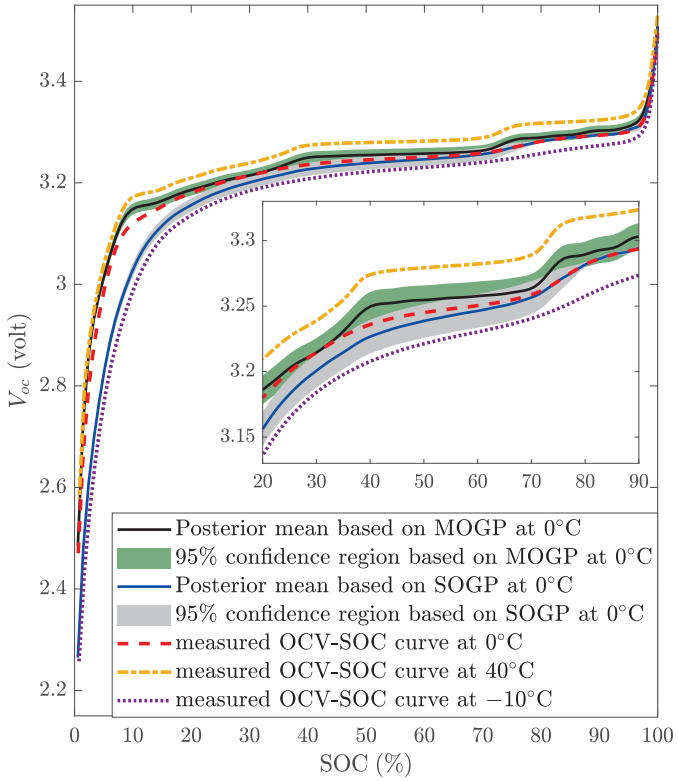

Figure 4: Identified OCV-SOC curve at $T_1 = 0°\text{C}$ using the proposed MOGP approach and SOGP approach in [24]

measurement noise are used as tuning parameters, which are denoted as $\bar{P}$, $\bar{Q}$, and $\bar{R}$, respectively. For the best estimation performance, these tuning parameters are tuned independently for each ECM model as follows:

1) Baseline ECM-ID:
   $\bar{P} = \text{diag}\left(\begin{bmatrix}1.1 \times 10^{-1} & 1 \times 10^{-3}\end{bmatrix}\right)$,
   $\bar{Q} = \text{diag}\left(\begin{bmatrix}1 \times 10^{-7} & 2.2 \times 10^{-3}\end{bmatrix}\right)$, $\bar{R} = 2.7 \times 10^{-2}$;
2) ECM-ID with non-updated OCV-SOC curve:
   $\bar{P} = \text{diag}\left(\begin{bmatrix}1 \times 10^{-1} & 1 \times 10^{-3}\end{bmatrix}\right)$,
   $\bar{Q} = \text{diag}\left(\begin{bmatrix}1 \times 10^{-11} & 1 \times 10^{-5}\end{bmatrix}\right)$, $\bar{R} = 1 \times 10^{-2}$;
3) ECM-ID with updated OCV-SOC curve:

$\bar{P} = \text{diag}\left(\begin{bmatrix}1 & 1\times10^{-4}\end{bmatrix}\right)$,
$\bar{Q} = \text{diag}\left(\begin{bmatrix}1\times10^{-10} & 4\times10^{-3}\end{bmatrix}\right)$, $\bar{R} = 5\times10^{-2}$.

The resulting RMSEs of SOC estimation are shown in Table 3. It can be seen that the ECM with updated OCV-SOC curve leads to a slightly higher RMSE than the baseline ECM, but much lower RMSE than the ECM with non-updated OCV-SOC curve, which is reduced by at least 14.0% under three working conditions. This illustrates the necessity of updating the OCV-SOC curve for accurate SOC estimation as the temperature varies.

Table 2: The identified ECM parameters R_s, R_1 and C_1 of three ECMs at $T_1 = 0°\text{C}$

	$R_s(\Omega)$	$R_1(\Omega)$	$C_1(\text{F})$
Baseline ECM-ID	0.196	0.068	296.236
ECM-ID with non-updated OCV-SOC curve	0.195	0.224	451.982
ECM-ID with updated OCV-SOC curve	0.190	0.024	330.837

Table 3: RMSEs of SOC estimation based three ECM models under three working conditions at $T_1 = 0°\text{C}$

	FUDS	US06	DST
Baseline ECM-ID	8.8×10^{-3}	4.9×10^{-3}	3.0×10^{-3}
ECM-ID with non-updated OCV-SOC curve	1.2×10^{-2}	1.3×10^{-2}	8.6×10^{-3}
ECM-ID with updated OCV-SOC curve	7.7×10^{-3}	5.4×10^{-3}	7.4×10^{-3}

6. Conclusions

The OCV-SOC curve varies with temperature, hence needs to be updated at a new temperature. In this paper, we propose a time-efficient OCV-SOC curve identification method using current-voltage data, without measuring or estimating OCVs. It fuses data from a piece of charging/discharging process at a given temperature as well as the OCV-SOC curve data at other temperatures. The proposed approach adopts an MOGP model to capture correlations among OCV-SOC curves at different temperatures, and predicts the OCVs at the given temperature by the posterior means. Using detailed comparisons with experimental data, we illustrate that the proposed MOGP approach achieves smaller errors in the OCV-SOC curve identification, and leads to more accurate SOC estimation.

Acknowledgment

This work is supported by the National Natural Science Foundation of China (Grant No. 61803163); Hubei Provincial Natural Science Foundation for Innovation Groups (Grant No. 2021CFA026); the Science and Technology Project of State Grid Corporation of China (Grant No. 5419-202199552A-0-5-ZN).

References

[1] S. M. Youssry, I. El-Hallag, R. Kumar, G. Kawamura, W. K. Tan, A. Matsuda, M. N. El-Nahass, Electrochemical deposition of uniform and porous Co–Ni layered double hydroxide nanosheets on nickel foam for supercapacitor electrode with improved electrochemical efficiency, Journal of Energy Storage 50 (2022) 104638.

[2] R. Kumar, S. Sahoo, E. Joanni, R. K. Singh, A review on the current research on microwave processing techniques applied to graphene-based supercapacitor electrodes: An emerging approach beyond conventional heating, Journal of Energy Chemistry (2022).

[3] E. Joanni, R. Kumar, W. P. Fernandes, R. Savu, A. Matsuda, In situ growth of laser-induced graphene micro-patterns on arbitrary substrates, Nanoscale 14 (25) (2022) 8914–8918.

[4] R. Kumar, E. Joanni, R. K. Singh, D. P. Singh, S. A. Moshkalev, Recent advances in the synthesis and modification of carbon-based 2d materials for application in energy conversion and storage, Progress in Energy and Combustion Science 67 (2018) 115–157.

[5] R. Kumar, R. K. Singh, D. P. Singh, E. Joanni, R. M. Yadav, S. A. Moshkalev, Laser-assisted synthesis, reduction and micro-patterning of graphene: recent progress and applications, Coordination Chemistry Reviews 342 (2017) 34–79.

[6] M. T. Lawder, B. Suthar, P. W. Northrop, S. De, C. M. Hoff, O. Leitermann, M. L. Crow, S. Santhanagopalan, V. R. Subramanian, Battery energy storage system and battery management system for grid-scale applications, Proceedings of IEEE 102 (6) (2014) 1014–1030.

[7] L. Lu, X. Han, J. Li, J. Hua, M. Ouyang, A review on the key issues for lithium-ion battery management in electric vehicles, Journal of Power Sources 226 (2013) 272–288.

[8] X. Chen, H. Lei, R. Xiong, W. Shen, R. Yang, A novel approach to reconstruct open circuit voltage for state of charge estimation of lithium ion batteries in electric vehicles, Applied Energy 255 (2019) 113758.

[9] R. Xiong, Q. Yu, L. Wang, C. Lin, A novel method to obtain the open circuit voltage for the state of charge of lithium ion batteries in electric vehicles by using H infinity filter, Applied Energy 207 (2017) 346–353.

[10] R. Zhang, B. Xia, B. Li, L. Cao, Y. Lai, W. Zheng, H. Wang, W. Wang, M. Wang, A study on the open circuit voltage and state of charge characterization of high capacity lithium-ion battery under different temperature, Energies 11 (9) (2018) 2408.

[11] R. Xiao, Y. Hu, X. Jia, G. Chen, A novel estimation of state of charge for the lithium-ion battery in electric vehicle without open circuit voltage experiment, Energy 243 (2022) 123072.

[12] L. Chen, X. Wu, A. M. Lopes, L. Yin, P. Li, Adaptive state-of-charge estimation of lithium-ion batteries based on square-root unscented kalman filter, Energy 252 (2022) 123972.

[13] C. Weng, J. Sun, H. Peng, A unified open-circuit-voltage model of lithium-ion batteries for state-of-charge estimation and state-of-health monitoring, Journal of Power Sources 258 (2014) 228–237.

[14] Y. Cui, P. Zuo, C. Du, Y. Gao, J. Yang, X. Cheng, Y. Ma, G. Yin, State of health diagnosis model for lithium ion batteries based on real-time impedance and open circuit voltage parameters identification method, Energy 144 (2018) 647–656.

[15] X. Bian, L. Liu, J. Yan, Z. Zou, R. Zhao, An open circuit voltage-based model for state-of-health estimation of lithium-ion batteries: Model development and validation, Journal of Power Sources 448 (2020) 227401.

[16] M. Zhang, G. Kang, L. Wu, Y. Guan, A method for capacity prediction of lithium-ion batteries under small sample conditions, Energy 238 (2022) 122094.

[17] A. Farmann, D. U. Sauer, A study on the dependency of the open-circuit voltage on temperature and actual aging state of lithium-ion batteries, Journal of Power Sources 347 (2017) 1–13.

[18] H. He, X. Zhang, R. Xiong, Y. Xu, H. Guo, Online model-based estimation of state-of-charge and open-circuit voltage of lithium-ion batteries in electric vehicles, Energy 39 (1) (2012) 310–318.

[19] F. Zheng, Y. Xing, J. Jiang, B. Sun, J. Kim, M. Pecht, Influence of different open circuit voltage tests on state of charge online estimation for lithium-ion batteries, Applied Energy 183 (2016) 513–525.

[20] Q. Zhang, N. Cui, Y. Li, B. Duan, C. Zhang, Fractional calculus based modeling of open circuit voltage of lithium-ion batteries for electric vehicles, Journal of Energy Storage 27 (2020) 100945.

[21] S. S. S. Narayanan, S. Thangavel, Machine learning-based model development for battery state of charge–open circuit voltage relationship using regression techniques, Journal of Energy Storage 49 (2022) 104098.

[22] L. Lavigne, J. Sabatier, J. M. Francisco, F. Guillemard, A. Noury, Lithium-ion Open Circuit Voltage (OCV) curve modelling and its ageing adjustment, Journal of Power Sources 324 (2016) 694–703.

[23] U. B. Vyas, V. A. Shah, Optimisation based 3-dimensional polynomial regression to represent lithium-ion battery's open circuit voltage as function of state of charge and temperature, Journal of Energy Storage 50 (2022) 104656.

[24] C. Huang, L. Wang, Gaussian process regression-based modelling of lithium-ion battery temperature-dependent open-circuit-voltage, Electronics Letters 53 (17) (2017) 1214–1216.

[25] A. Klintberg, C. Zou, B. Fridholm, T. Wik, Kalman filter for adaptive learning of two-dimensional look-up tables applied to OCV-curves for aged battery cells, Control Engineering Practice 84 (2019) 230–237.

[26] X. Tang, Y. Wang, C. Zou, K. Yao, Y. Xia, F. Gao, A novel framework for Lithium-ion battery modeling considering uncertainties of temperature and aging, Energy conversion and management 180 (2019) 162–170.

[27] C. Li, N. Cui, C. Wang, C. Zhang, Simplified electrochemical lithium-ion battery model with variable solid-phase diffusion and parameter identification over wide temperature range, Journal of Power Sources 497 (2021) 229900.

[28] M. Seo, Y. Song, J. Kim, S. W. Paek, G.-H. Kim, S. W. Kim, Innovative lumped-battery model for state of charge estimation of lithium-ion batteries under various ambient temperatures, Energy 226 (2021) 120301.

[29] Q.-K. Wang, Y.-J. He, J.-N. Shen, Z.-F. Ma, G.-B. Zhong, A unified modeling framework for lithium-ion batteries: An artificial neural network based thermal coupled equivalent circuit model approach, Energy 138 (2017) 118–132.

[30] M. Ouyang, G. Liu, L. Lu, J. Li, X. Han, Enhancing the estimation accuracy in low state-of-charge area: A novel onboard battery model through surface state of charge determination, Journal of Power Sources 270 (2014) 221–237.

[31] Y. Zheng, W. Gao, X. Han, M. Ouyang, L. Lu, D. Guo, An accurate parameters extraction method for a novel on-board battery model considering electrochemical properties, Journal of Energy Storage 24 (2019) 100745.

[32] K. Fan, Y. Wan, B. Jiang, State-of-charge dependent equivalent circuit model identification for batteries using sparse Gaussian process regression, Journal of Process Control 112 (2022) 1–11.

[33] H. Liu, J. Cai, Y.-S. Ong, Remarks on multi-output Gaussian process regression, Knowledge-Based Systems 144 (2018) 102–121.

[34] E. V. Bonilla, K. Chai, C. Williams, Multi-task Gaussian process prediction, Advances in Neural Information Processing Systems 20 (2007).

[35] C. K. Williams, C. E. Rasmussen, Gaussian Processes for Machine Learning, MIT press Cambridge, MA, 2006.

[36] A. Papoulis, P. S. Unnikrishna, Probability, Random Variables, and Stochastic Processes, McGraw-Hill, 2002.

[37] Center for Advanced Life Cycle Engineering at University of Maryland, Data description: A123, data retrieved at 2022 from `https://web.calce.umd.edu/batteries/data.htm#INR18650`.

[38] Y. Xing, W. He, M. Pecht, K. L. Tsui, State of charge estimation of lithium-ion batteries using the open-circuit voltage at various ambient temperatures, Applied Energy 113 (2014) 106–115.

[39] C. E. Rasmussen, H. Nickisch, The GPML toolbox version 4.2 (2022).

[40] C. Zhang, W. Allafi, Q. Dinh, P. Ascencio, J. Marco, Online estimation of battery equivalent circuit model parameters and state of charge using decoupled least squares technique, Energy 142 (2018) 678–688.