IEEE Control Systems Letters paper presented at
2023 American Control Conference (ACC)
San Diego, CA, USA. May 31 - June 2, 2023

Error Analysis for Parameter Estimation of Li-ion Battery subject to System Uncertainties

Jackson Fogelquist, Qingzhi Lai, and Xinfan Lin

Abstract— **Lithium-ion battery parameter estimation is a dynamic research field in which creative and novel algorithms are being developed to tune high-fidelity models for advanced control of energy systems. Amidst these efforts, little focus has been placed on the fundamental mechanisms associated with estimation accuracy, giving rise to the question, why is an estimate accurate or inaccurate? In response, we derive a generalized multivariate estimation error equation for the least-squares objective, which reveals that the error can be represented as the product of system uncertainties (i.e., in model, measurement, and parameter) and uncertainty-propagating sensitivity structures. We then relate the error equation to conventional error analysis criteria, such as parameter sensitivity, the Fisher information matrix, and the Cramér-Rao bound, to assess the benefits and limitations of each. Broadly, these criteria share the principal deficiency of neglecting estimation bias and system uncertainties, which are inevitable in practice. The error equation is validated through a series of experimental uni- and bivariate estimations of lithium-ion battery electrochemical parameters. These results are also analyzed using the error equation to study the composition of errors under various data sets. Finally, the bivariate analysis indicates that adding an additional target parameter to the estimation without increasing the amount of data intrinsically reduces the error robustness to the influence of system uncertainties.**

I. Introduction

Parameter estimation can be broadly defined as fitting a mathematical model to data through the identification of numerical constants. Accordingly, it is a vital element of model-based control, which strongly relies upon accurate models of physical systems. Recently, the accelerating deployment of electric vehicles and renewable energy systems has spurred significant research interest in the estimation of electrochemical parameters for lithium-ion (Li-ion) batteries [1], [2], [3], which enables the high-fidelity battery modeling capabilities necessary for advanced battery management systems [4] and degradation monitoring [5].

A parameter estimation problem is comprised of three elements—measurement data, system model, and estimation algorithm. The algorithm is typically a set of optimization procedures for determining the values of model parameters that minimize a cost function associated with the error between the measurement data and modeled output, e.g., sum of squared errors. As such, several challenges exist that negatively affect the estimation performance. For example, nonlinear models and large parameter sets can cause identifiability issues and high computational expense [3]. More notably, the measured data may not contain sufficient information about the target parameters, which can limit the attainable estimation accuracy [6].

Accordingly, estimation analysis methods have been applied to study the quality of data and its influence on the estimation result. In battery modeling and control, conventional analysis criteria are mostly based on parameter sensitivities, the Fisher information, and the Cramér-Rao bound. Sensitivity analysis examines how the variation of parameter values affect the output data. For example, [7] examined the sensitivity dynamics of electrochemical battery parameters through analytical sensitivity analyses, while [2] ranked the sensitivities of electrochemical battery parameters through numerical sensitivity analyses. Computed based on parameter sensitivities, the Fisher information can be used to quantify the amount of information about each parameter that is embedded in the data [8]. In [9], the eigenvalues of the Fisher information matrix were used to rank the identifiability of electrochemical battery parameters. The Fisher information matrix was analytically derived in [10] to examine the identifiability of a simple battery model under periodic excitation. Furthermore, the inverse of the Fisher information gives the Cramér-Rao bound, which characterizes the lower bound of the estimation error (co)variance for an unbiased estimator [8], [11]. In [12], [13], the Cramér-Rao bound was analytically derived for uni- and multivariate state and parameter estimation for a simple battery model under generic excitation, to quantify the role of battery characteristics and data on estimation accuracy. Additionally, since the Cramér-Rao bound indicates that error (co)variance is minimized when the Fisher information matrix is "maximized," scalar metrics of the Fisher information (e.g., determinant, trace, minimum eigenvalue) have become standard metrics for data quality and are typically implemented for optimal experiment design [14], [15], [16].

These existing estimation analysis criteria and methods are limited because they do not consider estimation bias nor system uncertainties. This significantly restricts their effectiveness, as estimation bias and system uncertainties are major sources of estimation error and are inevitable in practice [6]. Specifically, estimation accuracy is strongly influenced by constant and varying uncertainty in model (e.g., due to unmodeled dynamics) [17], measurement (e.g., due to sensor bias/noise) [18], and parameter (e.g., due to changing operating conditions/degradation) [19]. We sought to address these limitations in a prior work through the derivation of a univariate estimation error equation for the least-squares objective, which directly predicts the estima-

Jackson Fogelquist, Qingzhi Lai, and Xinfan Lin are with the Department of Mechanical and Aerospace Engineering, University of California, Davis, CA 95616, USA, corresponding author e-mail: lxflin@ucdavis.edu

tion error through consideration of uncertainties in model, measurement, and parameter [6]. This error equation was subsequently leveraged to develop criteria for data optimization in univariate estimation scenarios, i.e., cost functions for data design [6] and selection criteria for data selection [20], [21], which yielded promising results and outperformed the conventional Fisher information-based criteria.

The objective of this paper is to answer a fundamental question—why is an estimation result accurate or inaccurate?—both from a generic perspective and in the context of battery parameter estimation problems. The topic is investigated through the following contributions. First, we will derive and validate a generalized multivariate estimation error equation for the least-squares objective that is not subject to conventional limitations (i.e., unbiased estimation and omission of system uncertainties). This equation will reveal several important insights, such as the theoretical relationship between system uncertainties and estimation error, data structures that are capable of attenuating the influence of system uncertainties, and the precedence of data quality over quantity. Second, we will apply this equation to analyze the error composition in several battery electrochemical parameter estimation scenarios under various data sets. This will indicate the factors that enhance/degrade estimation accuracy and illustrate their strong dependence on data. Third, we will compare the errors and their compositions across uni- and bivariate estimation scenarios to identify an intrinsic mechanism that can reduce accuracy as additional parameters are simultaneously estimated under the same data. These contributions cast new light on the state of the art estimation error analysis criteria. Specifically, traditional local and global sensitivity analyses indicate that the magnitude of parameter sensitivity is important, yet we will show that the structure of the sensitivity can have an even greater impact on the estimation result. Similarly, Fisher information-based criteria specify that estimation performance will simply improve with the amount of data, though we will show that the structure of the data plays a more significant role and unfavorable data can potentially degrade the accuracy under uncertainty. Finally, the Cramér-Rao bound and other Fisher information-based criteria do not consider estimation bias or system uncertainties, though we will explicitly reveal how these inevitable factors are related to the estimation error.

II. Estimation Error Derivation & Analysis

In this section, we will derive and discuss the estimation error equation that will be implemented for error analysis in the proceeding sections. Consider a discrete-time single-input-single-output system model,

$$\begin{aligned} \boldsymbol{x}_k &= \boldsymbol{f}_k(\boldsymbol{x}_{k-1}, \boldsymbol{\theta}, \boldsymbol{\phi}, u_{k-1}) \\ y_k &= g_k(\boldsymbol{x}_k, \boldsymbol{\theta}, \boldsymbol{\phi}, u_k), \end{aligned} \tag{1}$$

where $\boldsymbol{x}$, u, and y are the state vector, input, and output of the system, $\boldsymbol{f}$ and g are the nonlinear state and output equations, and k is the time step index. The system is parameterized by $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$, where $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_n]^T$ is the vector of target parameters to be estimated and $\boldsymbol{\phi} = [\phi_1, \ldots, \phi_m]^T$ is the vector of non-estimated system parameters.

The objective of the estimation problem is to determine $\boldsymbol{\theta}$ based on a sequence of N measured output data $\boldsymbol{y}^m = [y_1^m, \ldots, y_N^m]^T$ sampled across consecutive time steps under input sequence $\boldsymbol{u} = [u_1, \ldots, u_N]^T$. To incorporate various system uncertainties, each output measurement y_k^m is expressed as

$$y_k^m = y_k(\boldsymbol{\theta}, \boldsymbol{\phi}, \boldsymbol{u}_k) + \Delta y + \delta y_k, \tag{2}$$

where $y_k(\boldsymbol{\theta}, \boldsymbol{\phi}, \boldsymbol{u}_k)$ is the modeled system output under the true parameter values $(\boldsymbol{\theta}, \boldsymbol{\phi})$, Δy is the constant bias between $y_k(\boldsymbol{\theta}, \boldsymbol{\phi}, \boldsymbol{u}_k)$ and the output measurement y_k^m due to measurement bias and/or model uncertainty, and δy_k is the varying model/measurement uncertainty due to factors such as sensor noise and unmodeled system dynamics. It is noted that $y_k(\boldsymbol{\theta}, \boldsymbol{\phi}, \boldsymbol{u}_k)$ denotes the mapping from $\boldsymbol{\theta}$, $\boldsymbol{\phi}$, and $\boldsymbol{u}_k$ to y_k with the state dynamics contained implicitly.

We employ the discrete-time multivariate least-squares objective—one of the most widely-used objectives for parameter estimation—to determine the estimated parameter set $\hat{\boldsymbol{\theta}}$ that minimizes the sum of squared errors between the measured output y_k^m and modeled output y_k across the N data points.

$$\min_{\hat{\boldsymbol{\theta}}} J = \frac{1}{2} \sum_{k=1}^{N} \left(y_k^m - y_k(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}}, \boldsymbol{u}_k) \right)^2 \tag{3}$$

It is noted that the estimation problem further includes parameter uncertainty in $\boldsymbol{\phi}$, as the exact values may not be known, which is indicated by the notation $\hat{\boldsymbol{\phi}}$ in Eqn. (3). We then apply the first-order optimality condition $\left(\nabla_{\hat{\boldsymbol{\theta}}} J = \boldsymbol{0}\right)$ to Eqn. (3), which yields

$$-\sum_{k=1}^{N} \frac{\partial y_k^T}{\partial \hat{\boldsymbol{\theta}}} (\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}}, \boldsymbol{u}_k) \left(y_k^m - y_k(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}}, \boldsymbol{u}_k) \right) = \boldsymbol{0}, \tag{4}$$

where $\frac{\partial y_k}{\partial \hat{\boldsymbol{\theta}}} = \left[\frac{\partial y_k}{\partial \hat{\theta}_1}, \ldots, \frac{\partial y_k}{\partial \hat{\theta}_n} \right]$ is the (row) vector of output sensitivity to each target parameter in $\hat{\boldsymbol{\theta}}$ at time step k, and $\boldsymbol{0}$ is the n-dimensional null column vector.

Recall that the measured output y_k^m is represented in function of the unknown true parameter values $(\boldsymbol{\theta}, \boldsymbol{\phi})$ in Eqn. (2). Therefore, we expand $y_k(\boldsymbol{\theta}, \boldsymbol{\phi}, \boldsymbol{u}_k)$ with a first-order Taylor series about the estimated/assumed parameter values $(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}})$, i.e.,

$$\begin{aligned} y_k(\boldsymbol{\theta}, \boldsymbol{\phi}, \boldsymbol{u}_k) \approx y_k(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}}, \boldsymbol{u}_k) &+ \frac{\partial y_k}{\partial \hat{\boldsymbol{\theta}}}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}}, \boldsymbol{u}_k) \Delta \boldsymbol{\theta} \\ &+ \frac{\partial y_k}{\partial \hat{\boldsymbol{\phi}}}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}}, \boldsymbol{u}_k) \Delta \boldsymbol{\phi}, \end{aligned} \tag{5}$$

where $\Delta \boldsymbol{\theta} = \boldsymbol{\theta} - \hat{\boldsymbol{\theta}}$ is the estimation error in $\hat{\boldsymbol{\theta}}$, $\Delta \boldsymbol{\phi} = \boldsymbol{\phi} - \hat{\boldsymbol{\phi}}$ is the parameter uncertainty in $\hat{\boldsymbol{\phi}}$, and $\frac{\partial y_k}{\partial \hat{\boldsymbol{\phi}}} = \left[\frac{\partial y_k}{\partial \hat{\phi}_1}, \ldots, \frac{\partial y_k}{\partial \hat{\phi}_m} \right]$ is the (row) vector of output sensitivity to each non-estimated parameter in $\hat{\boldsymbol{\phi}}$ at time step k.

Combining Eqns. (2), (4), & (5) and rearranging yields the multivariate estimation error equation in Eqn. (6), which expresses the estimation error $\Delta\boldsymbol{\theta}$ in terms of the parameter sensitivities and system uncertainties, i.e., the model/measurement bias Δy, varying model/measurement uncertainty δy_k, and parameter uncertainty vector $\Delta\boldsymbol{\phi}$. Note that the sensitivities are dependent on $\hat{\boldsymbol{\theta}}$, $\hat{\boldsymbol{\phi}}$, and $\boldsymbol{u}_k$, though these terms are omitted for brevity. If $\hat{\theta}$ and $\hat{\phi}$ are scalars, Eqn. (6) is reduced to the compact scalar form in Eqn. (7).

Several insights can be drawn from the error equation.

1) More data do not necessarily improve estimation accuracy. We illustrate this by examining the scalar form of the error equation in Eqn. (7), which indicates that, while increasing the number of data points (N) will monotonically increase the denominator, the numerator may also increase and possibly at a higher rate. For example, if the sensitivity $\frac{\partial y_k}{\partial \hat{\theta}}$ is small (<1), the approximately linear rate of increase of the numerator could outpace the quadratic rate of the denominator. Alternatively, consider that the varying model/measurement uncertainty δy_k may be very large during a portion of the data; the estimation error could be reduced by using a subset of the data that avoids the highly uncertain portion. This insight is corroborated by several works that have improved estimation accuracy through strategic data selection from a larger data set [20], [21], [22].
2) The term $\sum_{k=1}^{N} \frac{\partial y_k^T}{\partial \hat{\boldsymbol{\theta}}} \frac{\partial y_k}{\partial \hat{\boldsymbol{\theta}}}$ is essentially the Fisher information matrix simplified under independently and identically distributed (i.i.d.) Gaussian noise [10], [13]. Thus, we see that the Fisher information is directly related to the estimation error and serves as an approximate measure of robustness against the influence of system uncertainties. This is plainly illustrated in the scalar error equation in Eqn. (7), where a large Fisher information in the denominator will attenuate the impact of the system uncertainty terms in the numerator. However, the Fisher information is only part of the equation and is thus not a perfect standalone indicator of estimation accuracy. For example, Eqn. (7) shows that a large error will occur if the uncertainty terms in the numerator are significantly larger than the Fisher information in the denominator, regardless of the size of the Fisher information. The same conclusion can be obtained for the multivariate case by applying norm analysis to the general form in Eqn. (6). Therefore, it is critical to consider the role of system uncertainties in estimation error analysis. Additionally, we see that Fisher information-based criteria favor more data for estimation, which may not be optimal according to the previous insight on data length.
3) The term $\left(\sum_{k=1}^{N} \frac{\partial y_k^T}{\partial \hat{\boldsymbol{\theta}}}\right) \Delta y$ represents the error caused by the constant model/measurement uncertainty Δy. Since Δy is propagated to the total error through the sensitivity structure $\sum_{k=1}^{N} \frac{\partial y_k^T}{\partial \hat{\boldsymbol{\theta}}}$, the influence of Δy is eliminated if $\sum_{k=1}^{N} \frac{\partial y_k^T}{\partial \hat{\boldsymbol{\theta}}} = \mathbf{0}$, i.e., each target parameter sensitivity sums to zero across the data sequence. This relationship between a specific sensitivity structure and the estimation error illuminates the limitation of traditional sensitivity analysis methods, which only focus on the magnitude of sensitivity.
4) The term $\sum_{k=1}^{N} \frac{\partial y_k^T}{\partial \hat{\boldsymbol{\theta}}} \delta y_k$ represents the error caused by the varying model/measurement uncertainty δy_k. Essentially, this term is a vector of inner products between the sensitivity sequence vector for each target parameter $\frac{\partial \boldsymbol{y}}{\partial \hat{\theta}_i} = \left[\frac{\partial y_1}{\partial \hat{\theta}_i}, \ldots, \frac{\partial y_N}{\partial \hat{\theta}_i}\right]$ and the uncertainty sequence vector $\boldsymbol{\delta y} = [\delta y_1, \ldots, \delta y_N]^T$, where each vector consists of the respective entries at all time steps of the sequence. Accordingly, the influence of δy_k on the estimation error of a certain $\hat{\theta}_i$ is zero if $\sum_{k=1}^{N} \frac{\partial y_k^T}{\partial \hat{\theta}_i} \delta y_k = 0$, i.e., if the sensitivity sequence vector of $\hat{\theta}_i$ is orthogonal to the uncertainty sequence vector. This is an important sensitivity structure that propagates uncertainty to the estimation error, which is not considered by conventional analysis criteria.
5) The final term $\left(\sum_{k=1}^{N} \frac{\partial y_k^T}{\partial \hat{\boldsymbol{\theta}}} \frac{\partial y_k}{\partial \hat{\boldsymbol{\phi}}}\right) \Delta\boldsymbol{\phi}$ represents the error caused by uncertainty in the non-estimated model parameters $\boldsymbol{\phi}$. The associated sensitivity structure $\sum_{k=1}^{N} \frac{\partial y_k^T}{\partial \hat{\boldsymbol{\theta}}} \frac{\partial y_k}{\partial \hat{\boldsymbol{\phi}}}$ is a matrix of inner products between the sensitivity sequence vector of each estimated parameter $\hat{\theta}_i$ and that of each non-estimated parameter $\hat{\phi}_j$, i.e., $\frac{\partial \boldsymbol{y}}{\partial \hat{\theta}_i}$ and $\frac{\partial \boldsymbol{y}}{\partial \hat{\phi}_j}$, which is multiplied by the uncertainty vector $\Delta\boldsymbol{\phi}$. Thus, the influence of $\Delta\boldsymbol{\phi}$ on the total estimation error can be eliminated if $\sum_{k=1}^{N} \frac{\partial y_k^T}{\partial \hat{\boldsymbol{\theta}}} \frac{\partial y_k}{\partial \hat{\boldsymbol{\phi}}} = \mathbf{0}$, where each combination of $\frac{\partial \boldsymbol{y}}{\partial \hat{\theta}_i}$ and $\frac{\partial \boldsymbol{y}}{\partial \hat{\phi}_j}$ is orthogonal. This is an additional sensitivity structure with critical impact on estimation accuracy, which is related to our prior analytical Cramér-Rao bound analysis [13]. The study indicates that bivariate estimation does not suffer loss of accuracy (increased error covariance) over univariate estimation if the data contain orthogonal sensitivities between parameters.

III. Li-ion Battery Model & Sensitivity

This section will briefly summarize the Li-ion battery model, target parameters, and parameter sensitivity associated with our application of analyzing electrochemical parameter estimation errors. Specifically, we will investigate the estimation of the cathode lithium diffusion coefficient $D_{s,p}$, cathode active material volume fraction $\varepsilon_{s,p}$, and anode reaction rate constant k_n, based on the widely-adopted single particle model with electrolyte dynamics (SPMe) [7]. The involved parameters play a vital role in the performance of the battery and are commonly studied in state of health (SOH) estimation applications, as they indicate various degradation mechanisms [23]. The model and the sample parameters are used in this work as examples for illustrating the methodology of estimation error analysis, which can be applied to other models and parameters without loss of generality.

The SPMe captures the battery dynamics by predicting the output terminal voltage V and battery internal physical states from the input current I. The SPMe is a reduced-order model that represents the electrochemical processes in each

$$\Delta\boldsymbol{\theta} = -\left(\sum_{k=1}^{N} \frac{\partial y_k^T}{\partial \hat{\boldsymbol{\theta}}} \frac{\partial y_k}{\partial \hat{\boldsymbol{\theta}}}\right)^{-1} \left[\left(\sum_{k=1}^{N} \frac{\partial y_k^T}{\partial \hat{\boldsymbol{\theta}}}\right) \Delta y + \left(\sum_{k=1}^{N} \frac{\partial y_k^T}{\partial \hat{\boldsymbol{\theta}}} \delta y_k\right) + \left(\sum_{k=1}^{N} \frac{\partial y_k^T}{\partial \hat{\boldsymbol{\theta}}} \frac{\partial y_k}{\partial \hat{\boldsymbol{\phi}}}\right) \Delta \boldsymbol{\phi}\right] \tag{6}$$

$$\Delta\theta = -\frac{\left(\sum_{k=1}^{N} \frac{\partial y_k}{\partial \hat{\theta}}\right) \Delta y + \left(\sum_{k=1}^{N} \frac{\partial y_k}{\partial \hat{\theta}} \delta y_k\right) + \left(\sum_{k=1}^{N} \frac{\partial y_k}{\partial \hat{\theta}} \frac{\partial y_k}{\partial \hat{\phi}}\right) \Delta \phi}{\sum_{k=1}^{N} \left(\frac{\partial y_k}{\partial \hat{\theta}}\right)^2} \tag{7}$$

electrode (e.g., intercalation, diffusion) with a single particle, under the assumption that lithium intercalation current density is uniform across each electrode. Both electrode particles are interfaced through the electrolyte diffusion dynamics. The output terminal voltage is expressed as

$$\begin{aligned} V = &U_p(c_{se,p}) - U_n(c_{se,n}) + \phi_{e,p}(c_{e,p}) - \phi_{e,n}(c_{e,n}) \\ &+ \eta_p(c_{se,p}, c_{e,p}) - \eta_n(c_{se,n}, c_{e,n}) - IR_l, \end{aligned} \tag{8}$$

which includes the difference between the cathode and anode (denoted by subscripts p and n, respectively) in open-circuit potentials (OCPs) U, electrolyte potentials ϕ_e, and overpotentials η. These terms are functions of the dynamic lithium concentration states at the electrode particle surface (c_{se}) and electrolyte boundary (c_e), which are governed by Fick's second law of diffusion and the Butler-Volmer equation of (de)intercalation reaction kinetics at the particle surface. The voltage drop across various Ohmic resistances (e.g., SEI layer, current collectors) is incorporated through the lumped resistance term R_l. The reader is referred to [7] for the full details of the model.

The estimation error equation in Eqn. (6) relies upon the sensitivities of the estimated and non-estimated parameters. We employ the analytical sensitivity expressions derived for the SPMe in [7], which efficiently capture the dynamics of the sensitivity through sensitivity transfer functions (STFs). For example, the sensitivity of the output voltage V to $\varepsilon_{s,p}$ can be derived by taking the partial derivative of Eqn. (8) with respect to $\varepsilon_{s,p}$ as

$$\frac{\partial V}{\partial \varepsilon_{s,p}}(t) = \frac{\partial \eta_p}{\partial \varepsilon_{s,p}} + \left(\frac{\partial \eta_p}{\partial c_{se,p}} + \frac{\partial U_p}{\partial c_{se,p}}\right) \cdot \frac{\partial c_{se,p}}{\partial \varepsilon_{s,p}}(t). \tag{9}$$

In this equation, the first three terms can be easily computed based on the model, while $\frac{\partial c_{se,p}}{\partial \varepsilon_{s,p}}(t)$ is subject to dynamics, and needs to be computed based on the derived STF

$$\frac{\partial C_{se,p}}{\partial \varepsilon_{s,p}}(s) = \frac{7R_{s,p}^4 s^2 + 420 D_{s,p} R_{s,p}^2 s + 3465 D_{s,p}^2}{s(R_{s,p}^4 s^2 + 189 D_{s,p} R_{s,p}^2 s + 3465 D_{s,p}^2)} \frac{I(s)}{F \varepsilon_{s,p}^2 A_p \delta_p}, \tag{10}$$

e.g., by converting to a linear state-space representation. Here, A_p, δ_p, and $R_{s,p}$ are the cathode area, thickness, and particle radius, while F is the Faraday constant. The STF was derived based on the SPMe via Laplace transform and Padé approximation, with the full procedure and sensitivity expressions for other parameters detailed in [7].

IV. Univariate Estimation Error Analysis

We will first present the estimation error analysis in the univariate case. In the context of the battery parameter estimation problem, the output y is voltage V, and the input u is current I. Eqn. (6) will be applied to two univariate battery electrochemical parameter estimation problems with the purpose of (1) validating the error equation and (2) analyzing the composition of error sources due to different types of uncertainty. Both scenarios are subject to the varying model/measurement uncertainty δV_k, which characterizes the mismatch between the measured and modeled (under the true parameter values) outputs due to factors such as unmodeled system dynamics, discretization errors, and sensor bias/noise. Note that δV_k is the total varying model/measurement uncertainty, i.e., the sum of the constant (Δy) and varying (δy_k) components discussed in Section II. Both scenarios are also subject to parameter uncertainty $\Delta\boldsymbol{\phi}$, which characterizes the deviation between the true and assumed values of the non-estimated parameters due to factors such as manufacturing variation, degradation, and operating condition.

Each estimation is performed with the least-squares algorithm, SPMe battery model, and input-output data sets acquired through experimental testing of an LGM50T INR21700 Li-Nickel-Manganese-Cobalt (NMC) cell, using an Arbin LBT21084 cycler. Three dissimilar input current profiles were applied to the cell, namely, a constant 1C discharge (1C CC), 1C Pulse (1/60 Hz square wave), and the Federal Urban Driving Schedule drive-cycle (FUDS). Each profile has the same duration (30 minutes) and number of data points (6,000), and the initial cell state of charge (SOC) was 75% for the 1C CC profile and 50% for the 1C Pulse and FUDS profiles. The true parameter set was adapted from [24], which experimentally parameterized an LGM50 INR21700 cell through a full tear-down analysis.

The first scenario estimates the cathode active material volume fraction $\varepsilon_{s,p}$ subject to intrinsic model/measurement uncertainty and parameter uncertainty in the separator electrolyte porosity $\varepsilon_{e,sep}$ and cathode lithium diffusion coefficient $D_{s,p}$, i.e., $\theta = \varepsilon_{s,p}$ and $\boldsymbol{\phi} = [\varepsilon_{e,sep}, D_{s,p}]^T$. Parameter uncertainty was applied by scaling $\varepsilon_{e,sep}$ and $D_{s,p}$ by 20% from the true values. The general form of the error equation in Eqn. (6) is cast as Eqn. (11) for this scenario, where the denominator is the Fisher information of $\varepsilon_{s,p}$ and the numerator contains the uncertainty terms, δV_k, $\Delta\varepsilon_{e,sep}$, and $\Delta D_{s,p}$. The estimation was performed for each input current profile and the results are summarized in Table I. Specifically, the actual error is the true error between the estimate and the benchmark value, while the predicted error was computed from the error equation in Eqn. (11). The normalized Fisher information $\overline{F}_{info}$ is also included, which is the denominator of Eqn. (11) times $\hat{\varepsilon}_{s,p}^2$. The next three columns indicate the error contribution of each uncertainty, according to Eqn. (11). For example, the error contribution of $D_{s,p}$ is the third term in the numerator divided by the denominator. Finally, the last three columns list the root-mean-square (RMS) values of the model uncertainty δV_k and normalized sensitivities of the uncertain parameters, i.e., $\frac{\partial \overline{V}_k}{\partial \hat{\phi}} = \hat{\phi}\frac{\partial V_k}{\partial \hat{\phi}}$.

The second scenario estimates $D_{s,p}$ subject to intrinsic model/measurement uncertainty and parameter uncertainty in the anode lithium diffusion coefficient $D_{s,n}$ and $\varepsilon_{s,p}$. In this case, the assumed values of $D_{s,n}$ and $\varepsilon_{s,p}$ were selected to be 10% larger than the true values. The error equation for this scenario follows the same form as Eqn. (11), with $\varepsilon_{s,p}$, $\varepsilon_{e,sep}$, and $D_{s,p}$ replaced with $D_{s,p}$, $D_{s,n}$, and $\varepsilon_{s,p}$, respectively. The results are summarized in Table II.

Tables I & II indicate that the predicted errors agree well with the actual errors, validating the error equation in Eqn. (6). An important observation is that the deviations between the predicted and actual errors increase at larger actual errors. This is due to the first-order Taylor expansion in Eqn. (5), which was used to approximate the model output under the true parameter set $(\boldsymbol{\theta}, \boldsymbol{\phi})$ about the estimated/assumed parameter set $(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}})$. Since the accuracy of the expansion (and thus the error equation) is dependent on the proximity of the estimated/assumed parameter set to the true parameter set, prediction accuracy increases as the estimate of the target parameter approaches the true value, i.e., when the actual error is low. From a practical standpoint, we are typically not concerned with error prediction accuracy if the estimate is poor, yet even when this is the case, the error equation can still discern between high- and low-accuracy results, i.e., predicted errors are still large when actual errors are large, and small when actual errors are small.

A second observation is that the Fisher information is not necessarily correlated with the estimation error. For example, the 1C Pulse and FUDS results in Table I have the smallest and largest errors with nearly equivalent Fisher information values. In addition, the 1C CC Fisher information values in Tables I & II are one order of magnitude larger than those of FUDS, yet the estimation errors for 1C CC are higher. This highlights the limitations of the Fisher information as a standalone metric for data quality, though it is still intrinsic to the estimation error, as Eqn. (11) indicates that it is related to the robustness of the error against the uncertainty-propagating numerator terms. Even so, the system uncertainties play a significant role in the total error and should not be ignored.

Third, both Tables I & II reveal that the estimation error is strongly dependent on the data. This is because the model/measurement uncertainty and parameter sensitivities are dynamically driven by the input, e.g., as characterized by the sensitivity transfer functions discussed in Section III. It follows that the error contribution of each uncertainty is also dependent on the data, as observed by the variations across each table. For example, the dominant uncertainty in Table II is the model/measurement uncertainty δV_k under 1C CC and 1C Pulse, but the parameter uncertainty $\Delta\varepsilon_{s,p}$ is dominant under FUDS. In most cases, the uncertain parameter with the highest sensitivity has the largest error contribution, e.g., in Table II, $\varepsilon_{s,p}$ has a higher RMS sensitivity and error contribution than $D_{s,n}$ for every input profile. This aligns with the conventional sensitivity analysis criterion, which specifies that the most sensitive parameter will have the greatest influence on the estimation result. However, the influence of a parameter may not always be reflected by its sensitivity, as Table I reveals that $D_{s,p}$ is always more sensitive than $\varepsilon_{e,sep}$, though $\varepsilon_{e,sep}$ has a larger error contribution under 1C Pulse. This is due to the uncertainty-propagating sensitivity structures, e.g., $\sum_{k=1}^{N} \frac{\partial y_k}{\partial \hat{\theta}} \frac{\partial y_k}{\partial \hat{\phi}}$, which can amplify or attenuate the influence of an uncertain parameter based on the orthogonality of its sensitivity to that of the target parameter. Therefore, sensitivity magnitude is not sufficient as a standalone criterion for data quality—the uncertainty-propagating sensitivity structures are also critical.

V. Bivariate Estimation Error Analysis

We will now present the estimation error analysis in the bivariate case. The error equation in Eqn. (6) will be applied to a bivariate estimation scenario to validate the multivariate form of the error equation and analyze the effects of simultaneously estimating multiple parameters under the same data. The estimation procedure is identical to that of the univariate estimation in Section IV, except that two parameters are estimated instead of one, namely, $\varepsilon_{s,p}$ and the anode reaction rate constant k_n, giving $\boldsymbol{\theta} = [\varepsilon_{s,p}, k_n]^T$. Accordingly, the error equation in Eqn. (6) can be cast as Eqn. (12) for the error in $\varepsilon_{s,p}$, while the k_n error equation follows a symmetric form. The only uncertainty considered here is the model/measurement uncertainty δV_k, as the parameter uncertainty in $\varepsilon_{s,p}$ and k_n has effectively been eliminated by estimating both parameters simultaneously. When compared with the univariate form, the bivariate error equation in Eqn. (12) features a more complicated uncertainty-propagating sensitivity structure in the numerator and the determinant of the Fisher information matrix in the denominator.

The bivariate estimation results are summarized in Table III, where the predicted errors are presented alongside the actual errors for both parameters. The normalized determinant of the Fisher information matrix $\left|\overline{\boldsymbol{F}}_{info}\right|$ is also listed. For comparison, the univariate estimation results are included, where $\varepsilon_{s,p}$ and k_n were each independently estimated subject to model/measurement uncertainty and uncertainty in the respective non-estimated parameter, realized by a 20% deviation from the true value.

Table III first indicates that the actual and predicted errors mostly show good agreement for the bi- and univariate estimations, which further validates the error equation. However, exceptions occur in the bivariate results for $\varepsilon_{s,p}$ under 1C Pulse and k_n under 1C CC, which have significantly larger predicted errors than actual errors, i.e., 700% vs. 224% and 4004% vs. 724%, respectively. This is because the first-order Taylor expansion in Eqn. (5) causes error prediction accuracy to degrade when the actual error is large, as discussed in Section IV. Nevertheless, even at large actual errors, the error

$$\Delta\varepsilon_{s,p} = -\frac{\left(\sum_{k=1}^{N} \frac{\partial V_k}{\partial \hat{\varepsilon}_{s,p}} \delta V_k\right) + \left(\sum_{k=1}^{N} \frac{\partial V_k}{\partial \hat{\varepsilon}_{s,p}} \frac{\partial V_k}{\partial \hat{\varepsilon}_{e,sep}}\right) \Delta\varepsilon_{e,sep} + \left(\sum_{k=1}^{N} \frac{\partial V_k}{\partial \hat{\varepsilon}_{s,p}} \frac{\partial V_k}{\partial \hat{D}_{s,p}}\right) \Delta D_{s,p}}{\sum_{k=1}^{N} \left(\frac{\partial V_k}{\partial \hat{\varepsilon}_{s,p}}\right)^2} \tag{11}$$

TABLE I

UNIVARIATE ESTIMATION OF $\varepsilon_{s,p}$ SUBJECT TO UNCERTAINTY IN $\varepsilon_{e,sep}$ AND $D_{s,p}$

				Predicted Error Contribution			RMS Values (mV)		
Input Profile	Actual Error	Predicted Error	$\overline{F}_{info}$ (V^2)	δV_k	$\Delta\varepsilon_{e,sep}$	$\Delta D_{s,p}$	δV_k	$\frac{\partial \overline{V}_k}{\partial \hat{\varepsilon}_{e,sep}}$	$\frac{\partial \overline{V}_k}{\partial \hat{D}_{s,p}}$
1C CC	6.2%	7.5%	273	12.2%	-0.52%	-4.2%	38.8	6.8	55.3
1C Pulse	-22.3%	-17.3%	8.3	-17.0%	-0.73%	0.41%	21.5	2.0	4.3
FUDS	1.3%	2.1%	8.1	6.2%	-0.39%	-3.7%	8.4	1.0	8.6

TABLE II

UNIVARIATE ESTIMATION OF $D_{s,p}$ SUBJECT TO UNCERTAINTY IN $D_{s,n}$ AND $\varepsilon_{s,p}$

				Predicted Error Contribution			RMS Values (mV)		
Input Profile	Actual Error	Predicted Error	$\overline{F}_{info}$ (V^2)	δV_k	$\Delta D_{s,n}$	$\Delta\varepsilon_{s,p}$	δV_k	$\frac{\partial \overline{V}_k}{\partial \hat{D}_{s,n}}$	$\frac{\partial \overline{V}_k}{\partial \hat{\varepsilon}_{s,p}}$
1C CC	18.3%	25.7%	12.8	68.6%	-0.68%	-42.3%	38.8	5.2	221
1C Pulse	349%	463%	0.0082	529%	-0.47%	-65.9%	21.5	0.087	20.6
FUDS	-4.8%	-2.1%	0.60	29.9%	-0.081%	-31.9%	8.4	0.12	36.5

$$\Delta\varepsilon_{s,p} = \frac{\left(\sum_{k=1}^{N} \frac{\partial V_k}{\partial \hat{\varepsilon}_{s,p}} \frac{\partial V_k}{\partial \hat{k}_n}\right)\left(\sum_{k=1}^{N} \frac{\partial V_k}{\partial \hat{k}_n} \delta V_k\right) - \sum_{k=1}^{N} \left(\frac{\partial V_k}{\partial \hat{k}_n}\right)^2 \left(\sum_{k=1}^{N} \frac{\partial V_k}{\partial \hat{\varepsilon}_{s,p}} \delta V_k\right)}{\sum_{k=1}^{N} \left(\frac{\partial V_k}{\partial \hat{\varepsilon}_{s,p}}\right)^2 \sum_{k=1}^{N} \left(\frac{\partial V_k}{\partial \hat{k}_n}\right)^2 - \left(\sum_{k=1}^{N} \frac{\partial V_k}{\partial \hat{\varepsilon}_{s,p}} \frac{\partial V_k}{\partial \hat{k}_n}\right)^2} \tag{12}$$

equation is still effective at discerning the quality of data, i.e., large predicted errors are associated with large actual errors.

Second, the bivariate Fisher information does not solely determine the estimation error. For example, the bivariate Fisher information values under 1C CC and 1C Pulse are nearly equivalent, yet the $\varepsilon_{s,p}$ error is significantly larger under 1C Pulse, while the k_n error is larger under 1C CC. In addition, both univariate estimation cases under FUDS in Table III have the smallest errors despite having the smallest Fisher information values. These observations align with those of Section IV and reinforce that the Fisher information is limited as a standalone criterion for data quality, as the system uncertainties play a significant role.

Finally, the results in Table III reveal that simultaneously estimating two parameters can yield larger errors than estimating each parameter independently, under the same data. Specifically, we see that each bivariate estimation error is larger than the corresponding univariate estimation error, even though the former is not subject to parameter uncertainty through the inclusion of the unknown parameter in the estimation. This can be explained by examining the form of the $\varepsilon_{s,p}$ bivariate error equation in Eqn. (12). Specifically, dividing the numerator and denominator of Eqn. (12) by $\sum_{k=1}^{N}\left(\frac{\partial V_k}{\partial \hat{k}_n}\right)^2$ yields the new denominator, $\sum_{k=1}^{N}\left(\frac{\partial V_k}{\partial \hat{\varepsilon}_{s,p}}\right)^2 - \frac{\left(\sum_{k=1}^{N} \frac{\partial V_k}{\partial \hat{\varepsilon}_{s,p}} \frac{\partial V_k}{\partial \hat{k}_n}\right)^2}{\sum_{k=1}^{N}\left(\frac{\partial V_k}{\partial \hat{k}_n}\right)^2}$, which is the denominator of the univariate error equation in Eqn. (11) minus a positive term. Thus, the denominator always diminishes under the bivariate estimation case. Therefore, simultaneously estimating a second parameter causes bivariate estimation to be less robust than univariate estimation to the influence of system uncertainties, characterized by the numerator. This aligns with [12], [13], which show that the error covariance for multivariate unbiased estimation can never be less than that of univariate unbiased estimation, under the same data.

VI. CONCLUSIONS

In this paper, a generalized multivariate estimation error equation was derived for the least-squares objective that is not subject to the limitations of conventional analysis criteria, i.e., unbiased estimation and omission of system uncertainties. Analysis of the equation revealed the sensitivity structures that propagate various types of system uncertainties to the estimation error, and that estimation accuracy depends more on data quality than quantity. The error equation was validated through comparison of the predicted and actual errors in a series of uni- and bivariate battery electrochemical parameter estimation scenarios. These analyses highlighted the strong dependence of the estimation error (and its decomposition into the impact of different uncertainties) on data and the limitations of the conventional error analysis criteria. Specifically, parameter sensitivity magnitude alone does not determine the influence of uncertainties on the estimation result, as the uncertainty-propagating sensitivity structures are critical; the Fisher

TABLE III
BIVARIATE VERSUS UNIVARIATE ESTIMATION OF $\varepsilon_{s,p}$ AND k_n

Input Profile	Bivariate Estimation of $\varepsilon_{s,p}$ and k_n: $\varepsilon_{s,p}$ Act. Error	$\varepsilon_{s,p}$ Pred. Error	k_n Act. Error	k_n Pred. Error	$\lvert\overline{\boldsymbol{F}}_{info}\rvert$ (V^4)	Univariate Estimation of $\varepsilon_{s,p}$ Subject to Uncertainty in k_n: Act. Error	Pred. Error	$\overline{F}_{info}$ (V^2)	Univariate Estimation of k_n Subject to Uncertainty in $\varepsilon_{s,p}$: Act. Error	Pred. Error	$\overline{F}_{info}$ (V^2)
1C CC	-13.8%	-11.9%	724%	4004%	0.85	7.9%	8.8%	274	-21.5%	-9.5%	22.9
1C Pulse	224%	700%	-44.3%	-31.5%	0.66	-35.7%	-22.3%	16.6	-32.8%	-25.5%	32.3
FUDS	1.0%	1.0%	16.9%	19.0%	11.5	0.23%	1.0%	9.0	-0.011%	2.5%	3.1

information is an important term in the multivariate error equation associated with robustness to system uncertainties, yet it cannot serve as a standalone metric for data quality as the system uncertainties play a vital role; and the Cramér-Rao bound is not as widely applicable as the Fisher information as an error analysis criterion due to the highly restrictive assumption of unbiased estimation, as bias is inevitable in practice. Finally, the bivariate error analysis reinforced the insights from the univariate analysis and revealed that simultaneously estimating a second parameter under the same data intrinsically reduces the error robustness to system uncertainties. For further investigation, we refer the reader to our other work [25], which examines the error equation at greater depth with a mathematical correlation to the Cramér-Rao bound and a study on parameter identifiability. We have also been leveraging the identified uncertainty-propagating sensitivity structures for data design/selection to improve estimation accuracy [6], [20], [21].

ACKNOWLEDGMENT

We appreciate the funding support from the NSF CAREER Program (Grant No. 2046292) and the NASA HOME Space Technology Research Institute (Grant No. 80NSSC19K1052).

REFERENCES

[1] X. Lin, Y. Kim, S. Mohan, J. B. Siegel, and A. G. Stefanopoulou, “Modeling and estimation for advanced battery management,” *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 2, pp. 393–426, 2019.

[2] Z. Khalik, M. Donkers, J. Sturm, and H. Bergveld, “Parameter estimation of the Doyle-Fuller-Newman model for Lithium-ion batteries by parameter normalization, grouping, and sensitivity analysis,” *Journal of Power Sources*, vol. 499, p. 229901, Jul. 2021.

[3] W. Li, I. Demir, D. Cao, D. Jöst, F. Ringbeck, M. Junker, and D. U. Sauer, “Data-driven systematic parameter identification of an electrochemical model for lithium-ion batteries with artificial intelligence,” *Energy Storage Materials*, vol. 44, pp. 557–570, Jan. 2022.

[4] N. A. Chaturvedi, R. Klein, J. Christensen, J. Ahmed, and A. Kojic, “Algorithms for Advanced Battery-Management Systems,” *IEEE Control Systems*, vol. 30, no. 3, pp. 49–68, Jun. 2010.

[5] A. Allam and S. Onori, “On-line Capacity Estimation for Lithium-ion Battery Cells via an Electrochemical Model-based Adaptive Interconnected Observer,” *IEEE Transactions on Control Systems Technology*, vol. 29, no. 4, pp. 1636–1651, Jul. 2021, arXiv: 2008.10467.

[6] Q. Lai, H. J. Ahn, Y. Kim, Y. N. Kim, and X. Lin, “New data optimization framework for parameter estimation under uncertainties with application to lithium-ion battery,” *Applied Energy*, vol. 295, p. 117034, Aug. 2021.

[7] Q. Lai, S. Jangra, H. J. Ahn, G. Kim, W. T. Joe, and X. Lin, “Analytical derivation and analysis of parameter sensitivity for battery electrochemical dynamics,” *Journal of Power Sources*, vol. 472, pp. 228–338, 2020.

[8] T. M. Cover and J. A. Thomas, *Elements of Information Theory, 2nd Ed.* Hoboken, NJ: John Wiley & Sons, Inc., 2006, pp. 392–399.

[9] A. P. Schmidt, M. Bitzer, Ã. W. Imre, and L. Guzzella, “Experiment-driven electrochemical modeling and systematic parameterization for a lithium-ion battery cell,” *Journal of Power Sources*, vol. 195, no. 15, pp. 5071–5080, Aug. 2010.

[10] A. Sharma and H. K. Fathy, “Fisher identifiability analysis for a periodically-excited equivalent-circuit lithium-ion battery model,” in *2014 American Control Conference (ACC)*. Portland, OR, USA: IEEE, Jun. 2014, pp. 274–280.

[11] H. Cramér, *Mathematical Methods of Statistics (PMS-9)*. Princeton University Press, 1999.

[12] X. Lin and A. G. Stefanopoulou, “Analytic Bound on Accuracy of Battery State and Parameter Estimation,” *Journal of The Electrochemical Society*, vol. 162, no. 9, pp. A1879–A1891, 2015.

[13] X. Lin, “Analytic Analysis of the Data-Dependent Estimation Accuracy of Battery Equivalent Circuit Dynamics,” *IEEE Control Systems Letters*, vol. 1, no. 2, pp. 304–309, Oct. 2017.

[14] V. V. Fedorov, *Theory of Optimal Experiments*. New York, NY: Academic Press, Inc., 1972, pp. 27–30.

[15] S. Park, D. Kato, Z. Gima, R. Klein, and S. Moura, “Optimal experimental design for parameterization of an electrochemical lithium-ion battery model,” *Journal of The Electrochemical Society*, vol. 165, no. 7, pp. A1309–A1323, 2018.

[16] R. Huang, J. Fogelquist, and X. Lin, “Reinforcement learning of optimal input excitation for parameter estimation with application to Li-ion battery,” *IEEE Transactions on Industrial Informatics*, 2023.

[17] P. P. Mishra, M. Garg, S. Mendoza, J. Liu, C. D. Rahn, and H. K. Fathy, “How Does Model Reduction Affect Lithium-Ion Battery State of Charge Estimation Errors? Theory and Experiments,” *Journal of The Electrochemical Society*, vol. 164, no. 2, pp. A237–A251, 2017.

[18] X. Lin, “Theoretical analysis of battery SOC estimation errors under sensor bias and variance,” *IEEE Transactions on Industrial Electronics*, vol. 65, pp. 7138–7148, 2018.

[19] X. Lin, A. Stefanopoulou, P. Laskowsky, J. Freudenberg, Y. Li, and R. D. Anderson, “State of Charge Estimation Error due to Parameter Mismatch in a Generalized Explicit Lithium Ion Battery Model,” in *Dynamic Systems and Control Conference*, vol. 1, 10 2011, pp. 393–400.

[20] J. Fogelquist and X. Lin, “Uncertainty-aware data selection framework for parameter estimation with application to Li-ion battery,” in *2022 American Control Conference (ACC)*. IEEE, 2022, pp. 384–391.

[21] ——, “Data selection framework for battery state of health related parameter estimation under system uncertainties,” *eTransportation*, under review.

[22] X. Lin, “A data selection strategy for real-time estimation of battery parameters,” in *2018 Annual American Control Conference (ACC)*, 2018, pp. 2276–2281.

[23] V. Ramadesigan, K. Chen, N. A. Burns, V. Boovaragavan, R. D. Braatz, and V. R. Subramanian, “Parameter Estimation and Capacity Fade Analysis of Lithium-Ion Batteries Using Reformulated Models,” *Journal of The Electrochemical Society*, vol. 158, no. 9, p. A1048, 2011.

[24] C.-H. Chen, F. Brosa Planella, K. O'Regan, D. Gastol, W. D. Widanage, and E. Kendrick, “Development of Experimental Techniques for Parameterization of Multi-scale Lithium-ion Battery Models,” *Journal of The Electrochemical Society*, vol. 167, no. 8, p. 080534, Jan. 2020.

[25] J. Fogelquist, Q. Lai, and X. Lin, “On the error of Li-ion battery parameter estimation subject to system uncertainties,” *Journal of The Electrochemical Society*, 2023.