Journal of Power Sources 499 (2021) 229901

Contents lists available at ScienceDirect

Journal of Power Sources

journal homepage: www.elsevier.com/locate/jpowsour

Parameter estimation of the Doyle–Fuller–Newman model for Lithium-ion batteries by parameter normalization, grouping, and sensitivity analysis

Z. Khalik [a,*], M.C.F. Donkers [a], J. Sturm [c], H.J. Bergveld [a,b]

[a] Department of Electrical Engineering, Eindhoven University of Technology, Groene Loper 19, 5612AP Eindhoven, The Netherlands
[b] NXP Semiconductors, High Tech Campus 46, 5656 AE Eindhoven, The Netherlands
[c] Institute for Electrical Energy Storage Technology (EES), Technical University of Munich, Arcisstrasse 21, 80333, Munich, Germany

HIGHLIGHTS

- Parameters obtained from input/output data are not necessarily physically meaningful.
- Estimating all DFN model parameters is not necessary to obtain an accurate model.
- The length of estimation data should be carefully selected to avoid overfitting.
- Modeling errors can lead to a large bias and variability of the estimated parameters.
- Ideally, parameters should be determined using both cell teardown and estimation.

ARTICLE INFO

Keywords:
Battery parameter estimation
Doyle–Fuller–Newman model parameters
Doyle–Fuller–Newman model parameter estimation

ABSTRACT

Using electrochemistry-based battery models in battery management systems remains challenging due to the difficulty of uniquely determining all model parameters. This paper proposes a model parameterization approach of the Doyle–Fuller–Newman (DFN) model, by first reparameterizing the DFN model through normalization and grouping, followed by a sensitivity analysis and a parameter estimation procedure. In the parameter estimation procedure, we show the influence of the number of estimated parameters, as well as the influence of the data length of the identification data, on the obtained model accuracy. Additionally, the model with parameters obtained using the proposed parameterization approach is compared to a model whose parameters have been obtained using cell teardown. Finally, the consistency and accuracy of the parameter estimation procedure is analyzed by applying the estimation routine to a synthetic cell, represented by a DFN model with randomly chosen parameters. The results of this analysis show that the parameter estimation approach using current/voltage data can lead to a significantly better output accuracy, while it might not lead to physically meaningful parameters. This motivates the need for an approach that combines both and where cell tear-down can assist the parameter estimation using current/voltage data in achieving physically meaningful parameters.

1. Introduction

With the current trend in rising energy densities of Lithium-ion (Li-ion) batteries, battery-powered electric vehicles are becoming an increasingly competitive alternative to other common technologies, such as the combustion-engine vehicles. Battery models are required for battery-management-system functionalities, such as state-of-charge estimation [1], state-of-health estimation [2], and fast charging [3–6]. Due to their computational efficiency, equivalent-circuit models [7] are often used for this purpose, where the battery is modeled using (passive) circuit elements. However, these types of models provide limited information on the internal states of the battery, which can be important in control applications or in the analysis of battery behavior.

Therefore, electrochemistry-based Li-ion battery models have gained increasing popularity in such applications [2,3]. The Doyle–Fuller–Newman (DFN) model is a widely used electrochemistry-based model, which captures the physics on a macroscale (i.e., on a cell level) [8], using a set of partial differential equations (PDEs) [9], that describe the

* Corresponding author.
E-mail addresses: z.khalik@tue.nl (Z. Khalik), m.c.f.donkers@tue.nl (M.C.F. Donkers), johannes.sturm@tum.de (J. Sturm), h.j.bergveld@tue.nl (H.J. Bergveld).

https://doi.org/10.1016/j.jpowsour.2021.229901
Received 12 January 2021; Received in revised form 16 March 2021; Accepted 4 April 2021
Available online 3 May 2021
0378-7753/

internal states of the battery (consisting of solid-phase and electrolyte-phase Li concentrations and potentials). However, to accurately describe these internal states, many parameters need to be determined, which is a challenge, especially as determination of some parameter values requires complex experimental techniques and a highly time-consuming effort for the measurement procedure. Therefore, simplified electrochemistry-based models, such as the single-particle model (SPM) [10], have been proposed, which have fewer parameters. Consequently, much of the focus in model parameterization has been on such simplified models in the time domain [11–16] and on linearized models in the frequency domain [17]. However, the simplifications made in these models are usually at the cost of model accuracy, and therefore, parameterization of the DFN model, as has been done in, e.g., [18–27], is of interest. The interest in parameterizing the DFN model is also leveraged by the recent developments in highly efficient numerical implementations [28,29], which have led to a significant reduction in simulation time of the DFN model.

Generally, the DFN model parameters can be determined in several ways. One way is to determine several parameters based on information provided by the cell manufacturers, who know the material properties that have been chosen during cell design. However, the manufacturer often does not have all the parameter values or might not want to disclose them. Another way is to measure the parameters by cell teardown and experimental testing [23,24,26]. Such methods generally involve costly equipment, and it is sometimes not allowed to tear down a cell by the cell manufacturer. Finally, parameters can be determined through parameter estimation using input–output data [18–22,27]. In these methods, some, or all, of the parameters are estimated based on optimization of the model predictions to input/output measurements (i.e., external current/voltage measurements). Generally, two approaches are taken in the estimation of the parameters based on input/output measurements. One approach is to estimate some (or all) of the parameters simultaneously based on measured input/output data [18,19,21,27]. Since the identifiability of the DFN model is poor [19], a sensitivity analysis can be done to determine the parameters to which the model output is most sensitive, in order to select a smaller set of parameters for estimation [18]. The other approach is to design experiments specifically in an attempt to isolate the effects of parameters in the output [14,20,22]. Here, the parameters are grouped based on their effect to the output, after which the groups of parameters are estimated separately using their respectively designed input current. However, often, there is no justification given for this approach [14,20], or the approach is justified with the intuition that identifying too many parameters simultaneously may lead to unexpected uncertainty and errors [22]. However, this intuition has not yet been verified.

In this paper, we propose a model parameterization approach of the DFN model, by first reparameterizing the DFN model through normalization and grouping, followed by a sensitivity analysis of [30] that will show that some parameters have only a limited influence on the input–output behavior, and a parameter estimation procedure. The sensitivity analysis is similar to the approach taken in [18], in which the procedure of [30] is applied to a DFN model, but in which the parameters have not been grouped and normalized. The normalization and grouping procedure is similar to the approach taken in [25,31], in which no sensitivity analysis has been done. For the model parameterization, two cases are considered: one where the individual electrode potentials have been characterized, for which we have one set of experimental data available for parameterization, and one where only the electromotive force (EMF)/open-circuit potential (OCP) measurements of the complete cell are provided, but for which we have 3 experimental data sets available. In the parameter estimation routine, we will study the influence of the number of estimation parameters on the obtained model accuracy for both cases, as well as the influence of data length of identification data on accuracy. Finally, we will study the consistency and accuracy of the parameter estimation routine by analyzing the parameter estimation routine for a so-called synthetic cell, which is represented by a DFN model with a randomly chosen set of parameters. This will show how consistently and accurately the parameters of the DFN model can be estimated, and how far the resulting internal states deviate from the true internal states.

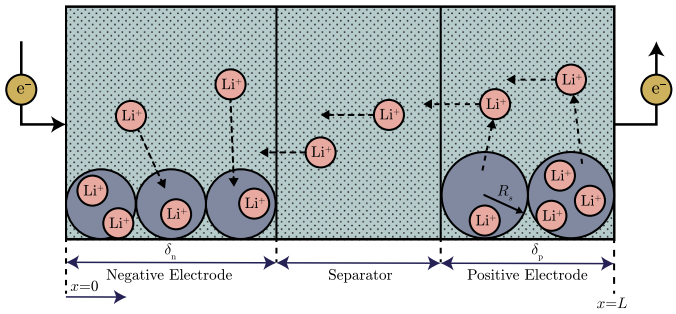

Fig. 1. DFN modeling approach for a Li-ion cell, where the depicted cell is being charged.

2. Battery modeling

The objective of this paper is to parameterize the DFN model using measured voltage and current data. This section first presents the DFN model, including its parameters and, subsequently, proposes to reparameterize the model to reduce the number of unknown parameters using an approach similar to that proposed in [25].

2.1. Doyle–Fuller–Newman model equations

The DFN model is a widely used electrochemistry-based model introduced in [9]. Fig. 1 illustrates the functionality of the model schematically for a Li-ion cell. The DFN model is often referred to as a pseudo-two-dimensional model, since the model describes concentrations and potentials in two dimensions, namely in a radial direction r along the radius of the particle, and in the x dimension, along the thickness of the electrode stack. In the x dimension, the cell is divided into three regions, namely the negative electrode, the separator, and the positive electrode. The reader is referred to [8] for a detailed description of the DFN model, and the underlying assumptions made.

The original model equations, i.e., before reparameterization, are based on the formulation given in [28], and references therein. These equations have been summarized in Table 1(a), given by (1)–(6), with the addition of a maximum reversible capacity constraint, given by (7). For compactness of notation, where possible, the time and space dependency of the variables have been left out of the equations. The symbols used in the original DFN model equations together with their respective descriptions are given in Table 2. Note that U in (5d) denotes the equilibrium potential of the electrode, which can be given by a pre-defined function typically of the solid-phase concentration at the solid-electrolyte interface ($c_{s,e}(x,t) = c_s(R_s,x,t)$). The solid-phase variables (i.e., $c_s(r,x,t)$, $\phi_s(x,t)$, $i_0(x,t)$, $U(x,t)$, and $\eta(x,t)$) are defined only in the electrodes, i.e., for $x \in [0,\delta_n] \cup [L-\delta_p, L]$, while the electrolyte-phase variables (i.e., $c_e(x,t)$ and $\phi_e(x,t)$), are defined over the whole length of the cell, i.e., for $x \in [0, L]$.

Over the three regions, the model parameters are defined as, e.g.,

$$\varepsilon_e(x) = \begin{cases} \varepsilon_{e,n} & \text{for } x \in [0,\delta_n] \\ \varepsilon_{e,\text{sep}} & \text{for } x \in [\delta_n, L-\delta_p] \\ \varepsilon_{e,p} & \text{for } x \in [L-\delta_p, L], \end{cases} \tag{15}$$

where the subscripts n, sep, and p refer to the negative electrode, separator, and positive electrode, respectively, and the other parameters are defined similarly. Because the parameters that only define a

Table 1
Model equations of the original and reparameterized DFN model.

	(a) Original model equations		(b) Reparameterized model equations	
Solid-phase Li-ion concentration	$\frac{\partial c_s}{\partial t} = \frac{D_s}{r^2}\frac{\partial}{\partial r}(r^2\frac{\partial c_s}{\partial r})$	(1a)	$\frac{\partial \hat{c}_s}{\partial t} = (\frac{\hat{D}_s}{\hat{r}^2})\frac{\partial}{\partial \hat{r}}(\hat{r}^2\frac{\partial \hat{c}_s}{\partial \hat{r}})$	(8a)
Boundary Condition	$\frac{\partial c_s}{\partial r}\Big\|_{r=0}=0,\ -D_s\frac{\partial c_s}{\partial r}\Big\|_{r=R_s}=j_n$	(1b)	$\frac{\partial \hat{c}_s}{\partial \hat{r}}\Big\|_{\hat{r}=0}=0,\ -\hat{D}_s\frac{\partial \hat{c}_s}{\partial \hat{r}}\Big\|_{\hat{r}=1}=\hat{j}_n$	(8b)
Electrolyte-phase Li-ion concentration	$\epsilon_e\frac{\partial c_e}{\partial t} = \frac{\partial}{\partial x}(D_e\epsilon_e^p\frac{\partial c_e}{\partial x})+\frac{3\epsilon_s(1-t_+^0)}{R_s}j_n$	(2a)	$\hat{\epsilon}_e\frac{\partial \hat{c}_e}{\partial t} = \frac{\partial}{\partial \hat{x}}(\hat{D}_e\hat{\rho}\frac{\partial \hat{c}_e}{\partial \hat{x}})+\hat{j}_n$	(9a)
Boundary condition	$\frac{\partial c_e}{\partial x}\Big\|_{x=0} = \frac{\partial c_e}{\partial x}\Big\|_{x=L} = 0$	(2b)	$\frac{\partial \hat{c}_e}{\partial \hat{x}}\Big\|_{\hat{x}=0} = \frac{\partial \hat{c}_e}{\partial \hat{x}}\Big\|_{\hat{x}=3} = 0$	(9b)
Solid-phase potential	$\frac{\partial}{\partial x}(\sigma\epsilon_s\frac{\partial \phi_s}{\partial x}) = \frac{3\epsilon_s F}{R_s}j_n$	(3a)	$\frac{\partial}{\partial \hat{x}}(\hat{\sigma}\frac{\partial \phi_s}{\partial \hat{x}}) = \hat{j}_n$	(10a)
Boundary condition	$\sigma\epsilon_s\frac{\partial \phi_s}{\partial x}\Big\|_{x=0}=\sigma\epsilon_s\frac{\partial \phi_s}{\partial x}\Big\|_{x=L}=\frac{i_{app}}{A}$	(3b)	$\hat{\sigma}\frac{\partial \phi_s}{\partial \hat{x}}\Big\|_{\hat{x}=0}=\hat{\sigma}\frac{\partial \phi_s}{\partial \hat{x}}\Big\|_{\hat{x}=3}=i_{\text{app}}$	(10b)
	$\frac{\partial \phi_s}{\partial x}\Big\|_{x=\delta_n} = \frac{\partial \phi_s}{\partial x}\Big\|_{x=L-\delta_p}=0$	(3c)	$\frac{\partial \phi_s}{\partial \hat{x}}\Big\|_{\hat{x}=1} = \frac{\partial \phi_s}{\partial \hat{x}}\Big\|_{\hat{x}=2}=0$	(10c)
Electrolyte-phase potential	$\frac{\partial}{\partial x}\left(\kappa\epsilon_e^p\frac{\partial \phi_e}{\partial x}+\kappa\epsilon_e^p(t_+^0\text{-}1)\frac{2RT}{F}\frac{\partial \ln c_e}{\partial x}\right)=-\frac{3\epsilon_s F}{R_s}j_n$	(4a)	$\frac{\partial}{\partial \hat{x}}\left(\hat{\kappa}\hat{\rho}\frac{\partial \phi_e}{\partial \hat{x}}+\hat{\kappa}\hat{\rho}(t_+^0-1)\frac{2RT}{F}\frac{\partial \ln \hat{c}_e}{\partial \hat{x}}\right)=-\hat{j}_n$	(11a)
Boundary condition	$\frac{\partial \phi_e}{\partial x}\Big\|_{x=0} = \phi_e\Big\|_{x=L} = 0$	(4b)	$\frac{\partial \phi_e}{\partial \hat{x}}\Big\|_{\hat{x}=0} = \phi_e\Big\|_{\hat{x}=3} = 0$	(11b)
Butler–Volmer kinetics	$j_n = \frac{i_0}{F}\left(\exp\left(\frac{\alpha_a F}{RT}\eta\right)-\exp\left(-\frac{\alpha_c F}{RT}\eta\right)\right)$	(5a)	$\hat{j}_n=\hat{i}_0\left(\exp\left(\frac{\alpha_a F}{RT}\eta\right)-\exp\left(-\frac{(1-\alpha_a)F}{RT}\eta\right)\right)$	(12a)
Exchange current density	$i_0 = k_0 c_e^{\alpha_a}(c_s^{\max}-c_{s,e})^{\alpha_a}c_{s,e}^{\alpha_c}$	(5b)	$\hat{i}_0 = \hat{k}_0\hat{c}_e^{\alpha_a}(\hat{c}_s^{\max}-\hat{c}_{s,e})^{\alpha_a}\hat{c}_s^{1-\alpha_a}$	(12b)
Particle surface concentration	$c_{s,e}(x,t) = c_s(R_s,x,t)$	(5c)	$\hat{c}_{s,e}(x,t) = \hat{c}_s(1,x,t)$	(12c)
Electrode over-potential	$\eta = \phi_s-\phi_e-U-FR_f j_n$	(5d)	$\eta = \phi_s-\phi_e-U-\hat{R}_f\hat{j}_n$	(12d)
Terminal voltage	$V(t) = \phi_s(L,t)-\phi_s(0,t)-\frac{R_{cc}}{A}i_{app}(t)$	(6)	$V(t) = \phi_s(3,t)-\phi_s(0,t)-\hat{R}_{cc}i_{app}(t)$	(13)
Maximum reversible capacity constraint	$Q = AF\delta_i\epsilon_{s,i}c_{s,i}^{\max}(s_{i,100\%}-s_{i,0\%}), i\in\{n,p\}$	(7)	$3Q = \hat{c}_{s,i}^{\max}(s_{i,100\%}-s_{i,0\%}), i\in\{n,p\}$	(14)

Table 2
List of symbols used in the original DFN model equations.

Symbol	Description	Unit
Latin		
A	Active electrode area	[m^2]
c_e	Electrolyte-phase Li-ion concentration	[mol/m^3]
$c_{e,0}$	Initial electrolyte concentration[a]	[mol/m^3]
c_s	Solid-phase Li-ion concentration	[mol/m^3]
$c_{s,e}$	Particle surface concentration	[mol/m^3]
$c_{s,n}^{\max}, c_{s,p}^{\max}$	Maximum solid-phase concentration[a]	[mol/m^3]
D_e	Li-ion diffusion coefficient in electrolyte[a]	[m^2]
$D_{s,n}, D_{s,p}$	Solid-phase Li-ion diffusion coefficient[a]	[m/s]
F	Faraday's constant	[C/mol]
i_0	Exchange current density	[A/m^2]
i_{app}	Applied current	[A]
j_n	Pore wall flux of Li-ions	[mol/m^2/s]
$k_{0,n}, k_{0,p}$	Kinetic constant[a]	[b]
L	Cell thickness[a]	[m]
p_n, p_{sep}, p_p	Bruggeman porosity exponent[a]	[–]
Q	Maximum reversible cell capacity	[C]
r	Radial position across a spherical particle	[m]
R	Universal gas constant	[J/mol/K]
R_{cc}	Current collector contact resistance[a]	[Ωm^2]
$R_{f,n}, R_{f,p}$	Particle surface film resistance [a]	[Ωm^2]
$R_{s,n}, R_{s,p}$	Radius of active material particles[a]	[m]
$s_{n,0\%}, s_{p,0\%}$	Stoichiometry at 0% state of charge[a]	[–]
$s_{n,100\%}, s_{p,100\%}$	Stoichiometry at 100% state of charge[a]	[–]
t	Time	[s]
t_+^0	Transference number[a]	[–]
U	Electrode equilibrium potential	[V]
V	Terminal voltage	[V]
x	Position across cell	[m]
Greek		
α_a	Anodic charge-transfer coefficient[a]	[–]
α_c	Cathodic charge-transfer coefficient[a]	[–]
δ_n, δ_p	Electrode thickness[a]	[m]
$\epsilon_{e,n}, \epsilon_{e,sep}, \epsilon_{e,p}$	Electrolyte volume fraction[a]	[–]
$\epsilon_{s,n}, \epsilon_{s,p}$	Active particles volume fraction[a]	[–]
η	Electrode over-potential	[V]
κ	Ionic conductivity[a]	[S/m]
σ_n, σ_p	Electrical conductivity[a]	[S/m]
ϕ_e	Electrolyte-phase potential	[V]
ϕ_s	Solid-phase potential	[V]

[a]Considered parameters of the DFN model.
[b][C/s · (m/mol)^(1 + 3α_c)]

property of the electrodes, e.g., D_s or σ, do not have a value defined in the separator region, i.e., for $x \in [\delta_n, L-\delta_p]$, the total number of parameters of the DFN model as formulated in Table 1(a) amounts to 35. Note that depending on how the DFN model is formulated, and what kind of assumptions are made, this number can vary. For example, most papers, e.g. [25,28,32], consider either the particle surface film resistance R_f or the current-collector resistance R_{cc} as a parameter to model any additional voltage drop that is generally observed at the output. However, as both these parameters can be linked to a physical phenomenon, we consider both of these effects, i.e., film resistance and current-collector resistance, as having separate parameters, in order to study the sensitivity of the output to these parameters.

2.2. Parameter normalization and grouping

From the equations of the DFN model in Table 1(a), it can be observed that the variation of certain parameters leads to the same physical effect. For example, by decreasing the diffusion coefficient D_s and increasing the particle size R_s accordingly, the diffusion dynamics in the solid phase remain unchanged. Thus, intuitively speaking, this means that one of these parameters is redundant, and the desired effect can instead be described with a single parameter. Mathematically, this can be achieved by normalization of (1), where the radial coordinate r is normalized through division by R_s. This results in a new radial coordinate $\hat{r} = r/R_s \in [0,1]$, after which the parameter R_s can be absorbed in a new parameter $\hat{D}_s = D_s/R_s^2$. Similarly, in the other Eqs. (2)–(4), the coordinate x can be redefined to

$$\hat{x} = \begin{cases} \frac{x}{\delta_n}, & 0 \le x \le \delta_n \\ 1+\frac{x-\delta_n}{L-\delta_p-\delta_n}, & \delta_n \le x \le L-\delta_p \\ 2+\frac{x-L+\delta_p}{\delta_p}, & L-\delta_p \le x \le L, \end{cases} \tag{16}$$

and the parameters can be grouped. This results in a new set of normalized Eqs. (8)–(14), shown in Table 1(b), with a new set of parameters resulting from the grouping of the original set of parameters, shown in Table 3. In the process, some of the variables have also been redefined, i.e.,

$$\hat{j}_n = \frac{3\epsilon_s FA\delta_i}{R_s}j_n, \quad \hat{c}_e = \frac{c_e}{c_{e,o}}, \quad \frac{\hat{c}_s}{3} = \epsilon_s FA\delta_i c_s, \tag{17}$$

Table 3
Parameters of the reparameterized DFN model.

Parameter	Grouping	Range	Unit
Q	Q	N/A	[C]
$s_{n,0\%}$	$s_{n,0\%}$	[0.002, 0.04]	[–]
$s_{p,0\%}$	$s_{p,0\%}$	[0.86, 0.97]	[–]
$s_{n,100\%}$	$s_{n,100\%}$	[0.75, 0.89]	[–]
$s_{p,100\%}$	$s_{p,100\%}$	[0.22, 0.44]	[–]
$\hat{D}_{s,n}$	$D_{s,n}/R_{s,n}^2$	[0.00013, 0.0016]	[s^{-1}]
$\hat{D}_{s,p}$	$D_{s,p}/R_{s,p}^2$	[0.0004, 0.63]	[s^{-1}]
$\hat{D}_e$	$D_e FAc_{e,0}/(1-t_+^0)$	$[4.1\cdot10^{-7}, 7.9\cdot10^{-6}]Q$[a]	[Cs^{-1}]
$\hat{p}_n$	$\varepsilon_{e,n}^{p_n}/\delta_n$	[50, 5700]	[–]
$\hat{p}_p$	$\varepsilon_{e,p}^{p_p}/\delta_p$	[58, 4100]	[–]
$\hat{p}_{sep}$	$\varepsilon_{e,sep}^{p_{sep}}/(L\text{-}\delta_n\text{-}\delta_p)$	[3700, 31000]	[–]
t_+^0	t_+^0	[0.26, 0.38]	[–]
$\hat{\sigma}_n$	$\sigma_n\varepsilon_{s,n}A/\delta_n$	$[1.2, 170]Q$[a]	[Ω^{-1}]
$\hat{\sigma}_p$	$\sigma_p\varepsilon_{s,p}A/\delta_p$	$[0.011, 8.3]Q$[a]	[Ω^{-1}]
$\hat{\kappa}$	κA	$[7\cdot10^{-6}, 2\cdot10^{-5}]Q$[a]	[Ω^{-1}]
$\hat{R}_{f,n}$	$R_{f,n}R_{s,n}/(3A\delta_n\varepsilon_{s,n})$	$[20, 330]/Q$[a]	[Ω]
$\hat{R}_{f,p}$	$R_{f,p}R_{s,p}/(3A\delta_p\varepsilon_{s,p})$	$[0, 0]/Q$[a]	[Ω]
$\hat{R}_{cc}$	R_{cc}/A	$[32, 170]/Q$[a]	[Ω]
α_a	α_a	[0.48, 0.52]	[–]
$\hat{k}_{0,n}$	$k_{0,n}c_{e,o}^{\alpha_a}/(R_{s,n}F)$	$[5.7\cdot10^{-5}, 0.00078]$	[b]
$\hat{k}_{0,p}$	$k_{0,p}c_{e,o}^{\alpha_a}/(R_{s,p}F)$	$[7.9\cdot10^{-5}, 0.001]$	[b]
$\hat{\varepsilon}_{e,n}$	$\varepsilon_{e,n}FA\delta_n c_{e,o}/(1-t_+^0)$	$[0.017, 0.76]Q$[a]	[C]
$\hat{\varepsilon}_{e,p}$	$\varepsilon_{e,p}FA\delta_p c_{e,o}/(1-t_+^0)$	$[0.01, 0.14]Q$[a]	[C]
$\hat{\varepsilon}_{e,sep}$	$\varepsilon_{e,sep}FA(L\text{-}\delta_n\text{-}\delta_p)c_{e,o}/(1-t_+^0)$	$[0.0049, 0.083]Q$[a]	[C]

[a]Note that some parameter ranges are scaled with the cell capacity Q.
[b]$[s^{-1}\cdot(m/mol)\hat{}(3(1-2\alpha_a))]$

where $i = n$ for $x \in [0, \delta_n], i = p$ for $x \in [L-\delta_p, L]$.

Note that, although the approach taken here is similar to the one in [25], there are a few differences. The differences mainly originate from the formulation of the original DFN model. In [25], the current collector resistance R_{cc} is not accounted for, and the diffusion and conductivity coefficients are formulated as their effective counterparts, e.g., $D_e^{\text{eff}} = D_e\varepsilon_e^p$, from which D_e^{eff} is chosen as the original model parameter. By treating the non-effective diffusion and conductivity coefficients as the model parameters, and basing the reparameterization procedure on this parameter set, an effectively lower amount of parameters is obtained here than in [25]. Furthermore, by considering the maximum reversible capacity constraint (7), an additional parameter is reduced, as $c_{s,n}^{\max}$ and $c_{s,p}^{\max}$ can be determined from the reversible capacity Q using this constraint. The total number of parameters of the reparameterized model is 24. While the total number of parameters of the reparameterized model presented in [25] is also 24, it has to be considered that we have an additional parameter (R_{cc}), and in [25] a further assumption is made that $\alpha_a = \alpha_c = 0.5$, which leads to reduction of an additional parameter. Thus, the reparameterized model proposed here has effectively two parameters less.

3. Model parameterization approach

This section describes a systematic parameterization approach to identify the parameters of the reparameterized model. First, an equilibrium potential model, which is used to identify the stoichiometric values and maximum concentrations, is described. Then the steps to identify the remaining parameters are explained, which include defining new ranges, the sensitivity analysis, and the parameter estimation procedure.

3.1. Equilibrium potential model

Besides the determination of the 24 parameters listed in Table 3, the equilibrium potential curves U_n and U_p need to be characterized. The equilibrium potential curves give the so-called EMF (or the OCP) of the battery as

$$U_{\text{EMF}}(s_c) = U_p(s_p) - U_n(s_n), \tag{18a}$$

where s_c refers to the state-of-charge (SOC) of the cell, and s_p and s_n, refer to the stoichiometry at the negative and positive electrode, respectively, defined as

$$s_i = \frac{\hat{c}_s}{\hat{c}_{s,i}^{\max}}, \tag{18b}$$

where $i = n$ for $\hat{x} \in [0, 1], i = p$ for $\hat{x} \in [2, 3]$. The electrodes of the battery generally have a larger capacity than the reversible capacity of the cell, and therefore s_n and s_p generally do not cycle between 0 and 1. Instead, when the battery is empty, i.e., at 0% SOC ($s_c = 0$), s_n and s_p have a certain pre-defined value (related to the balancing of the cell) of $s_{n,0\%}$ and $s_{p,0\%}$, respectively. Similarly, at 100% SOC ($s_c = 1$), s_n and s_p have a certain pre-defined value of $s_{n,100\%}$ and $s_{p,100\%}$, respectively. These parameters, together with U_n, U_p, and the maximum reversible cell capacity Q comprise the equilibrium potential model. Note that while the same EMF-SOC relation can be reached with any choice between 0 and 1 for the parameter values of $s_{n,0\%}$, $s_{n,100\%}$, $s_{p,0\%}$, $s_{n,100\%}$, these parameters also affect the dynamics of the cell through the $(\hat{c}_{s,\max} - \hat{c}_{s,e})^{\alpha_a}$ term in (12b). Therefore, these parameters should be considered in the parameter estimation routine, rather than assuming their values from literature to complete the equilibrium potential model, as done in, e.g., [18,33].

Generally, to determine the equilibrium electrode potential curves of the cell, the EMF of the cell needs to be determined as a function of SOC from input/output data. There are several techniques available to determine the EMF, where an overview can be found in [34]. From these measurements, the maximum reversible battery capacity Q can also be determined. However, in principle, the equilibrium potential curves of the electrodes cannot be determined individually from input/output data, and different techniques are required depending on whether a cell teardown can be performed or not. We therefore consider the following two cases:

(1) The electrode equilibrium potential curves cannot be determined through cell teardown. In this case, the electrode equilibrium potential curves can be determined using (18a), by assuming the negative equilibrium potential U_n from literature, from which U_p can be determined, as done in e.g. [16,18,33]. Since the negative electrode is generally the same or similar (usually a graphite-based composite) across different types of battery chemistries, and the negative electrode equilibrium potential U_n is relatively small compared to U_p, we propose to take U_n from literature. In this approach, it is necessary to choose the values of $s_{n,0\%}$, $s_{n,100\%}$, $s_{p,0\%}$, $s_{n,100\%}$, before U_p can be determined.
(2) The electrode equilibrium potentials can be measured through cell teardown. In this case, the equilibrium potential curves are measured in relation to the stoichiometry of the electrode over a sufficiently wide range of the stoichiometry. Then, the stoichiometric values $s_{n,0\%}$, $s_{n,100\%}$, $s_{p,0\%}$, $s_{n,100\%}$ are chosen such that the difference between the measured $U_p - U_n$ and the measured U_{EMF} is minimized in some way [23,24,26]. For further details on how the stoichiometric values are chosen in this case, we refer to the aforementioned papers.

Note that in the first case, information on the chemistry of the electrode materials is not required. If the chemistry of the electrode materials is known, but the electrode equilibrium potentials cannot be measured through cell teardown, then it is possible to determine an equilibrium potential model by taking the approach described for the second case, by taking the equilibrium potential curves from a database or literature, as done (or described) in, e.g., [14,15,20,21,25]. However, since the EMF can vary even between individual cells that were made in the same factory, the obtained EMF using this method may not sufficiently describe the measured EMF of the considered cell, especially if the purpose is to estimate some or all of the remaining DFN model parameters from input/output data. Deviations of the modeled EMF from the measured EMF can significantly impact the identifiability

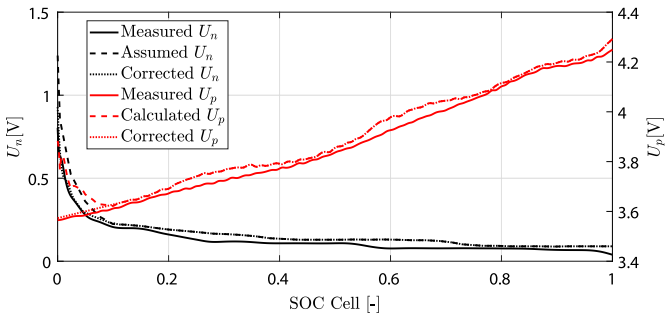

Fig. 2. Illustration of the correction of the calculated equilibrium potentials. Here, the measured equilibrium potentials are taken from [24], and the assumed negative electrode equilibrium potential curve is taken from [23].

of the model parameters, as also shown in [13]. Therefore, even if the chemistry of the electrode materials is known, it may still be preferable to use the approach described in the first case, since in that approach it is ensured that the modeled EMF coincides exactly with the measured EMF (by definition).

We should further note that the obtained U_p using the approach described in the first case often leads to a physically non-intuitive function for U_p. Fig. 2 demonstrates this, where the measured equilibrium potentials from [24] are compared to the calculated equilibrium potentials. Here, the assumed U_n has been obtained from [23], from which U_p is calculated from the EMF function shown in Fig. 3 using the method described in the first case above. We can observe that around 10% SOC, the calculated U_p rises significantly, which is contrary to what would be physically intuitive, as is demonstrated by the measured U_p. To correct for this, we replace the values of U_p from 0% to 10% SOC with a linear function, where the resulting corrected U_p is shown in Fig. 2 with the dotted lines. Here, the slope is chosen as the mean of the slope of the calculated U_p curve between 10% and 40% SOC. Note that to ensure that the modeled EMF still matches the measured EMF, U_n must also be corrected, which can be computed from the measured EMF and the corrected U_p as $U_n = U_p - U_{\text{EMF}}$. This corrected U_n is also shown in Fig. 2, where we can see that the corrected U_n also more closely matches the measured U_n, compared to the initially assumed U_n.

3.2. Ranges of the grouped parameters

After the parameters of the equilibrium potential model have been determined, the rest of the parameters can be determined through a parameter estimation routine, similar to what has been done in [18] for a model without normalized and grouped parameters. This involves performing a sensitivity analysis, and estimating (some of the) parameters from input/output data. The sensitivity analysis requires a range of values that the estimation parameters can have in order to assess the sensitivity of each parameter. Furthermore, the ranges can be used as bounds in the parameter estimation routine.

The ranges of the reparameterized model parameters are shown in Table 3. These parameter ranges have been obtained by gathering a set of values for each of the parameters from various papers, i.e., [23,24,26,35–37], where largely actual measurements have been done to obtain the parameter values (rather than using parameter estimation techniques to determine the parameters). Since the parameters given in these papers are for the non-reformulated DFN model, the parameters need to be converted to the parameters of the reformulated model, which can be done using Table 3. Note that in the computation of the reformulated DFN model parameter values, many of these parameters have a grouping that contains any of the parameters A, L, $\epsilon_{s,n}$, and $\epsilon_{s,p}$, which are related to the capacity of the cell. In order to account for this, the parameters that have any of the capacity-related parameters in their grouping are scaled by the capacity of the cell, as can be seen in Table 3. Further note that the range of $\hat{R}_{f,p}$ is $[0, 0]$, since we could not find any non-zero values for $\hat{R}_{f,p}$ in the prior-mentioned literature that we have used to determine the parameter ranges.

Within the given ranges, every parameter can be expressed on a linear or a logarithmic scale, i.e,

$$\theta_i = \beta_i \underline{\theta}_i + (1 - \beta_i)\overline{\theta}_i, \tag{19a}$$

$$\log \theta_i = \beta_i \log \underline{\theta}_i + (1 - \beta_i) \log \overline{\theta}_i, \tag{19b}$$

where $\beta_i \in \beta \in [0, 1]$, in which β denotes the vector of normalized parameters, $\theta_i \in \theta$ denotes every parameter, in which θ denotes the vector of all parameters, and $\underline{\theta}_i$, $\overline{\theta}_i$ denotes the minimum and maximum value of the parameter, respectively. For parameters whose range is in the same order of magnitude, i.e., $\overline{\theta}_i/\underline{\theta}_i \leq 10$, (19a) is used and for parameters whose range is of a different order of magnitude, i.e., $\overline{\theta}_i/\underline{\theta}_i > 10$, (19b) is used.

3.3. Sensitivity analysis of the grouped parameters

With the parameter ranges determined, a sensitivity analysis can be performed in order to determine the sensitivity of the output to the variation of the parameters. The results of the analysis can be used to select the appropriate (most sensitive) estimation parameters, in order to avoid solving an ill-conditioned optimization problem in the parameter estimation routine. The sensitivity analysis is done as described in [18], which in turn is based on the work of [30], where a parameter sensitivity ranking is obtained by orthogonalization of the sensitivity matrix of a model. Therefore, we will shortly summarize the sensitivity analysis approach, and refer to [30] and [18] for a more detailed description.

If we define the model output $\hat{V}$ as a nonlinear function g of the parameters $\theta(\beta)$, input i_{app}, and initial conditions x_0, i.e., $\hat{V} = g(\theta(\beta), i_{\text{app}}, x_0)$, the sensitivity of the model output $\hat{V}$ to the parameters θ can be defined as

$$S(\theta(\beta), i_{\text{app}}, x_0) = \frac{\partial \hat{V}}{\partial \beta} \tag{20}$$

where the gradient $\frac{\partial \hat{V}}{\partial \beta} \in \mathbb{R}^{n \times m}$, in which n is the number of time samples and m is the number of parameters, is numerically approximated using a finite-difference method. A permutation matrix $P \in \mathbb{R}^{m \times m}$ to determine the ranking of the sensitivity can be obtained using the pivoted QR decomposition of S such that

$$SP = QR, \tag{21}$$

where $Q \in \mathbb{R}^{n \times n}$ is an orthogonal matrix, and $R \in \mathbb{R}^{n \times m}$ is an upper-triangular matrix. If $\Pi_{\text{init}} = [1, 2, \ldots, m]$ denotes the initial ordering of the parameters, the set of rank-ordered parameters Π_{ranked} can be obtained through $\Pi_{\text{ranked}} = \Pi_{\text{init}} P$. Information on the magnitude of sensitivity M_s of the parameters is contained in the matrix R, and the magnitudes are given by the absolute value of the values on the diagonal of R, i.e., $M_s = |\text{diag}(R)|$. Note that the sensitivity matrix as defined in (20), and therefore the parameter ranking Π_{ranked}, depend on the chosen input profile as well as the initial conditions. For the purpose of selecting the appropriate estimation parameters, the input profile and initial conditions can be chosen the same as in the parameter estimation routine.

3.4. Estimating the grouped parameters

After performing the sensitivity analysis and determining the ranking order, the selected estimation parameters β_s can be optimized to the model output by minimizing the sum of the squared error between the experimental voltage V_{exp} and the predicted voltage $\hat{V}$, i.e.,

$$\hat{\beta}_s := \arg\min_{\beta} \sum_{i=1}^{n} (V_{\text{exp}}(t_i) - \hat{V}(\theta(\beta), t_i))^2, \tag{22}$$

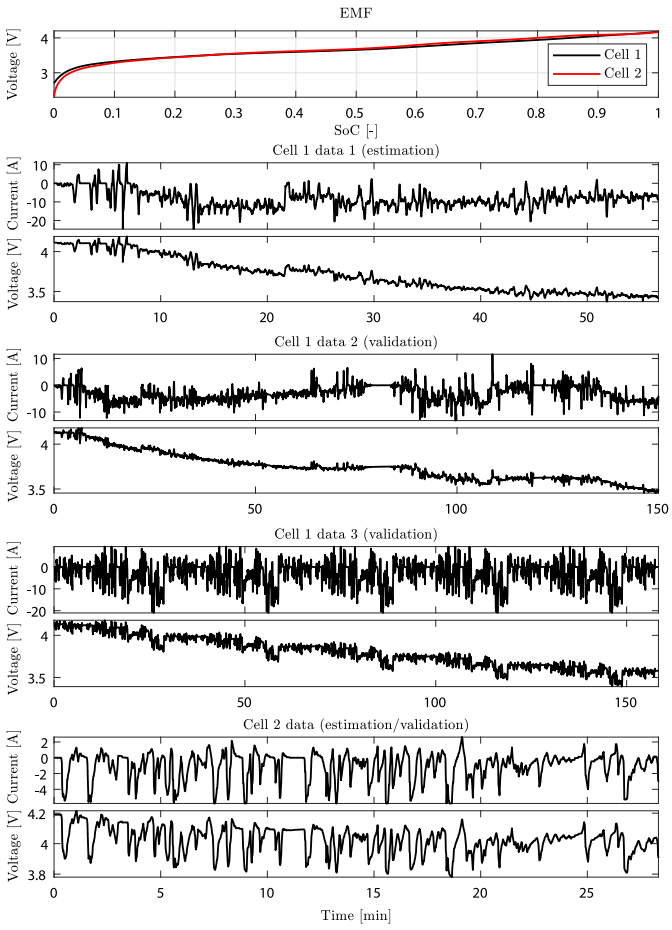

Fig. 3. Experimental data from Cell 1 and Cell 2 used for model parameterization.

where $\hat{\beta}_s$ is the vector of optimized parameters, and t_i denotes the time at which measurement $i \in \{1, \ldots, n\}$ is taken. This optimization problem can be solved using any nonlinear least-squares algorithm. The parameters that are not very sensitive will not be optimized and are selected as the nominal value in their ranges, i.e., by setting $\beta_i = 0.5$ in (18).

4. Results and validation

In this section, we will use the model parameterization approach presented in the previous section for two different cells. For one cell, only the EMF (or OCP) data of the complete cell is provided (Cell 1, see [38]), but 3 sets of experimentally obtained current and voltage measurements are available, and for the other cell, the equilibrium potential functions of the individual electrodes have been provided (Cell 2, see [24,39]), but for which we only have one set of current and voltage measurements available. We will first show the parameter sensitivity ranking for the two cells, then investigate the influence of the number of estimation parameters and the data length of identification data on obtained model accuracy. Furthermore, for Cell 2, a DFN model has already been parameterized in [24], where some of the parameters have been determined experimentally, and others have been assumed as values obtained from literature. We will compare the model accuracy obtained using the parameters determined from [24] and the model accuracy obtained using the model parameterization approach described in Section 3, and show that a significant improvement can be made by using the presented model parameterization approach. Finally, we will analyze the consistency and accuracy of the parameter estimation routine using a synthetic cell, represented by a DFN model of which its parameters have been randomly selected. The estimated parameters, and the resulting internal states will be compared to the true known parameters and states when the model used for estimation is equal to the true system, and when there are modeling errors present. This will show how consistently and accurately the parameters of the DFN model can be estimated, and how far the resulting internal states deviate from the true internal states. Note that measurements cannot be used here, since measured parameters and states are not available from current–voltage measurements.

The experimental data sets used for the sensitivity analysis and parameter estimation are shown in Fig. 3. For Cell 1, three different data sets were available, where one is used for the estimation of the parameters, and the other two are used for validation of the obtained parameters. However, for Cell 2, only one data set was available, which is used for both estimation as well as validation. Furthermore, for Cell 2, the equilibrium potentials have been measured, which are shown in Fig. 2.

4.1. Sensitivity analysis

In order to show how to make a selection of parameters to be estimated in the parameter estimation routine, the sensitivity analysis described in Section 3.3 has been applied. The current input profiles for the sensitivity matrix S in (20) have been chosen as the estimation profiles shown in Fig. 3. The parameter sensitivity ranking Π_{ranked} and the normalized magnitude of sensitivity (to the most sensitive parameter) $M_s/\max(M_s)$ of each parameter obtained for the two cells can be seen in Fig. 4. Note that since we assume that the reversible capacity Q is measurable and because the range of $\hat{R}_{f,p}$ in Table 3 was determined to be $[0, 0]$, the remaining number of parameters to be estimated is 22. We can observe that, while the normalized magnitudes and the order of the ranking are slightly different, the parameters occur roughly in the same order for both cells. Based on these rankings, for any desired number of estimation parameters, a selection of parameters can be made. These results have also been used in the studies shown below.

4.2. Influence of the number of estimated parameters on the accuracy

With a choice of the number of estimation parameters, the parameter estimation routine described in Section 3.4 has been done using the estimation current profiles shown in Fig. 3 for both cells. The nonlinear least-squares problem formulated in (22) has been solved using the `lsqnonlin` function in MATLAB with the trust-region-reflective algorithm, in which the parameter bounds have been selected as those listed in Table 3, and the initial parameter guesses have been obtained by setting $\beta = 0.5$ for the estimation parameters. For Cell 1, the model with the estimated parameters can be validated using a different data set than the one used for estimation. Note that based on the results shown in Fig. 5(b), which will be further explained below, only 90% of Data 1 has been used for estimation. For Cell 2, as only one dataset was available, the first 70% of the data has been used for estimation, and the entire data has been used for validation.

Fig. 5(a) shows the variation in root-mean-square error (RMSE) between the various data of Fig. 3 and the model with a varying number of estimated parameters. Here, the selected estimation parameters have been chosen in the order that corresponds to the ranking shown in Fig. 4. For both cells, we can see that after about 12 parameters the RMSE for all data sets does not significantly change. This seems to indicate that 12 estimation parameters is a fair selection in order to reach a small RMSE. For both cells, the RMSE seems to fluctuate somewhat at some points, e.g., for Cell 2, the validation RMSE with 14 estimation parameters is lower than with 15 estimation parameters, even though the estimation RMSE does not show this fluctuation. This can be explained by the fact that the parameter estimation problem is non-convex, and therefore there are multiple local minima, as we will also show further below.

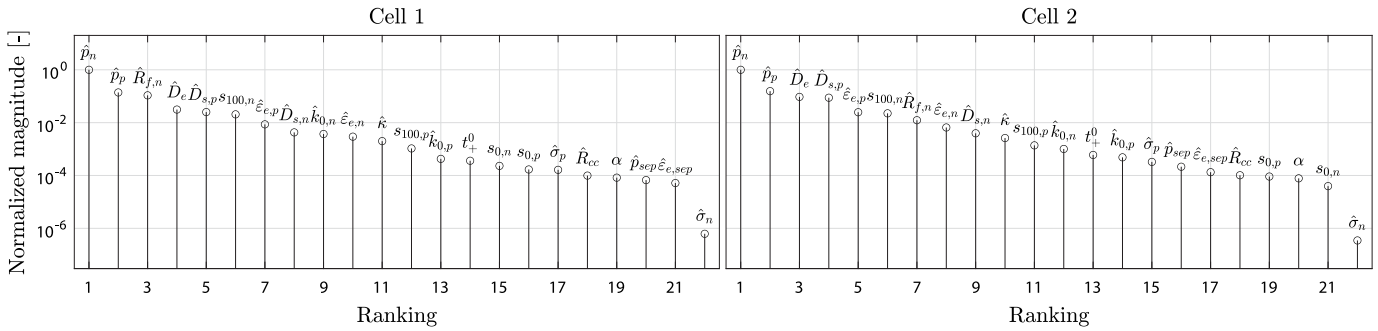

Fig. 4. Normalized magnitude of sensitivity for Cell 1 and Cell 2, where the parameters are ranked by decreasing magnitude of sensitivity.

4.3. Influence of the length of estimation data

The lack of a second set of data for Cell 2 means that the estimated model can only be validated using the original data used for estimation. Of course, validating a model with the exact same data as the model has been estimated with, does not give much confidence that this model will also accurately predict the output using a different current profile. Therefore, one way is to only use a subset of the available data for estimation, and then using the entire data for validation, as we have done in the previous subsection, where 70% of the data was used for estimation. However, the choice for the length of the estimation data is not obvious, as having less estimation data can lead to a worse estimate, while having less validation data can lead to false conclusions. In order make this apparent trade-off between a better model fit with more estimation data and a higher confidence in validation with more validation data, in Fig. 5(b), the RMSE of the validation for both cells is shown for a varying percentage of data used for estimation. Note that for Cell 1, the percentage is varied for the estimation data, and validated for all available data. We can observe that for both cells, the RMSE varies considerably based on the percentage of data used. It seems that for both cells, choosing at least around 70% of the estimation data is required to obtain a good model fit, which leaves 30% of the data to be validated over if there is only one data set available. Also notable is that for Cell 1, the RMSE is relatively small for all data sets even when only using 10% of the data. This is in contrast to Cell 2, where the obtained errors until using 65% of the data are much larger. This could be caused by more unmodeled behavior showing up in the experimental voltage of Cell 2 than of Cell 1.

In the lower plot of Fig. 6, we can see a comparison of the errors obtained using 30% of the data, 65% of the data, and with the parameters from [24]. We can see that when using 30% of the data for estimation (corresponding to the first 510 s), the errors are relatively small until 800 s, after which the error increases considerably again. When using 65% of the data (corresponding to 1105 s), the error remains small, for the entire data. This could indicate that between 800 s and 1105 s there are some significantly large unmodeled effects leading to large errors when less than 65% of the data is used. Furthermore, we can see that the simulation using the parameters from [24] shows a clear deviation from the experimental voltage, even larger than when using 30% of the data for estimation. Thus, we can see that by estimating some of the parameters, rather than relying on experiments and values from literature, a significantly more accurate model output can be obtained. However, we may still question whether the internal states represented by the model obtained from estimating the parameters are still physically meaningful. This will be discussed in the remaining subsections below.

Furthermore in Fig. 5(b), for Cell 1, we can see that with 100% of data used for estimation, there seems to be a sharp rise in the RMSE of Data 2 and Data 3, whereas the RMSE of Data 1 decreases considerably compared to using 90% of the estimation data. This indicates that there is a considerable case of increased overfitting of the parameters when using 100% of the estimation data. If we observe the upper plot in Fig. 6, we can see that for Data 1 there are significantly larger errors at the beginning, between 0.2 and 0.25 SoC, which explains the large increase in RMSE seen from 90% length of estimation data used in Fig. 5(b). These observations show that it is crucial to study where the unmodeled effects in the measured output occur the most, in order to avoid overfitting of the parameters to this unmodeled behavior as much as possible.

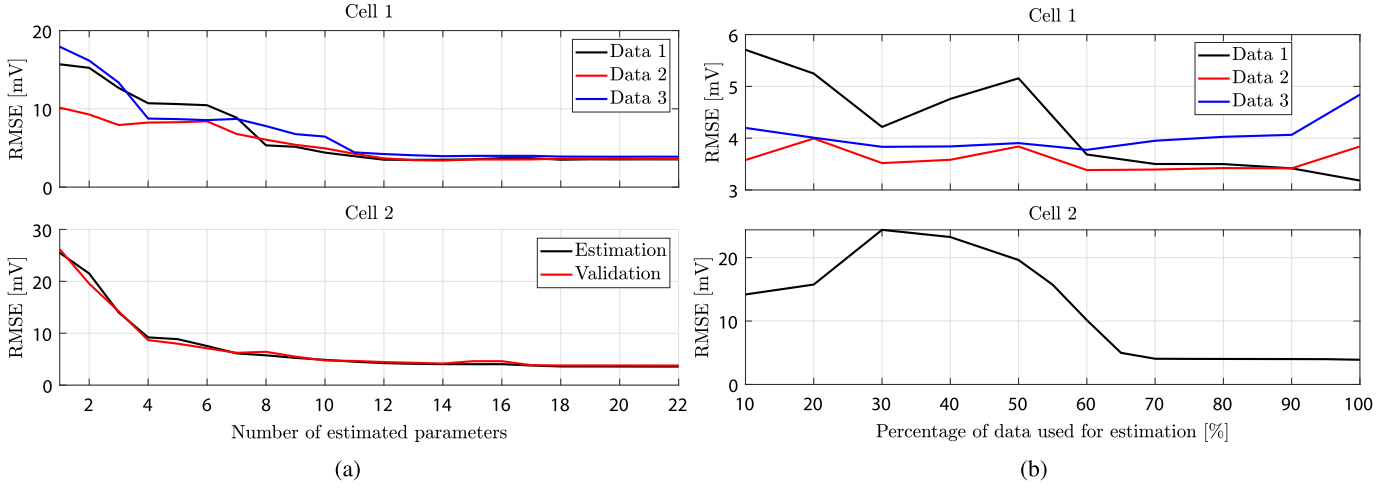

Fig. 5. RMSE between the measured data and simulated data for (a) a varying number of estimated parameters used in the estimation procedure and (b) a varying data length used in the estimation procedure, where for Cell 1, the percentage data used from Data 1 has been varied.

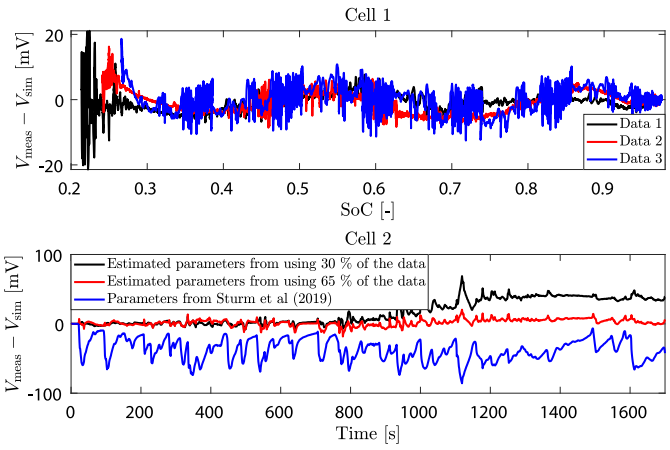

Fig. 6. Error between the measured voltage V_{meas} and the simulated voltage V_{sim} for both cells and their available data. For Cell 2, the error from using a DFN model with the parameters from [24] is also shown. The parameters used for simulation for Cell 1 have been obtained from using 60% of Data 1 for simulation with 14 estimated parameters and the parameters used for simulation for Cell 2 have been obtained from using 30% and 65% of the data for simulation with 14 estimated parameters.

4.4. Consistency and accuracy of the estimated parameters

The goal of parameter estimation is to find model parameters through fitting the modeled output and the measured output. If the identifiability of the parameters is sufficiently large, the estimated parameters should correspond to physically meaningful parameters. However, identifiability has been a key issue for the DFN model [19], which means that not all parameters can be found reliably. Furthermore, the nonlinear nature of the DFN model also means that there are multiple local minima, further complicating the issue. Finally, due to modeling errors, it can occur that even when the (internal) dynamics of the actual system is actually closer to a particular DFN model, a larger error in the validation of the terminal voltage is seen. Consequently, a smaller output validation error from a particular DFN model does not necessarily imply a better representation of the internal states of the system. In order to get an idea of how much these issues affect the parameter estimation procedure, we can perform the procedure from several random initial conditions and study the obtained parameters. Specifically, for Cell 1, we have done the parameter estimation routine, with 22 estimation parameters and 90% of Data 1 used for estimation, for 50 different random initial conditions of the parameter estimation routine.

The results of this study are shown in the upper plot of Fig. 7, where a boxplot representation is used to show the obtained data for each of the parameters. We can see here that it is also clear that most parameters can be estimated consistently. However, it also seems that many of the estimated parameters are at the extremal values, e.g., parameter 6 ($s_{100,n}$) and 8 ($\hat{D}_{s,n}$). This could be caused either by the unmodeled behavior in the experimental voltage, to which the parameters are being fit, or it could be that the ranges chosen for the parameters are too small. Therefore, we have done the same study, but this time with an increased range for the parameters, where the lower bound is found by setting $\beta_i = -0.3$ in (18) and the upper bound is found by setting $\beta_i = 1.3$ in (18). We can now see that, as could be expected, the variability of the estimated parameters increases. The variability is especially more for the lower ranked parameters, which is expected, as these were also found to be the least sensitive parameters according to the sensitivity analysis done above. For example, parameter 6 ($s_{100,n}$), which was consistently estimated as $\beta = 1$, now varies between around 0.2 and 1.3 (where we should point out that with $\beta = 1.3$, the parameter value is still physically meaningful, i.e., $s_{100,n} < 1$), although the median is still around $\beta = 1$. It could be somewhat unexpected that

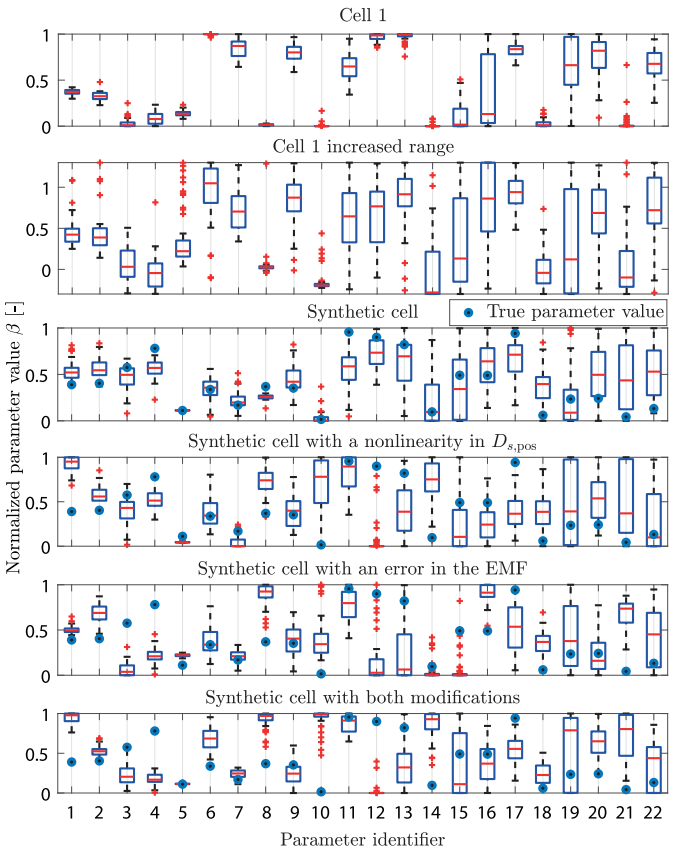

Fig. 7. Boxplots showing the variability and accuracy of the parameter estimation routine for Cell 1 and the synthetic cell. The parameter identifiers relate to the ranking shown in Fig. 4.

the parameter after extending the range also has estimated values that are below $\beta = 1$, since with the original ranges it did not go below $\beta = 1$. However, because the parameters have some dependency on one another, by extending the ranges, the number of combinations of parameters that achieves a local minimum increases. This can explain why, for parameter 6 with the extended ranges, we get combinations within the original range that would not show up with the original ranges.

Of course, it could also be that the physically meaningful or "true" parameters of the cell are in fact in the original ranges, but that the parameters have been fit so much to the unmodeled behavior that they tend to go to the extremes of the ranges. In order to investigate the impact of unmodeled behavior on the estimation of the parameters, we have generated a synthetic cell, of which its parameters have been chosen randomly (within the specified ranges). In this case, the estimation model is now exactly the same as the system that is to be estimated, which is considered to be the ideal scenario for parameter estimation. With this modeled cell, we can do the same study as outlined above (with the same input data used for simulation), where we start from 50 different random initial conditions.

The results of this simulation study are shown in the third plot of Fig. 7. We can now see that the variability of the parameters is considerably more than for Cell 1 in Fig. 7. This can be explained by the fact that for the synthetic cell, the values of the parameters fall within their respective parameter ranges (as this is how the parameters have been selected), while with Cell 1, this is not guaranteed, where the "true" parameter values can be out of their respectively chosen parameter ranges. Generally, as we have seen from the discussion of Fig. 7, when one parameter shows a higher variability, it will also lead to a higher variability in the other parameters, since there is

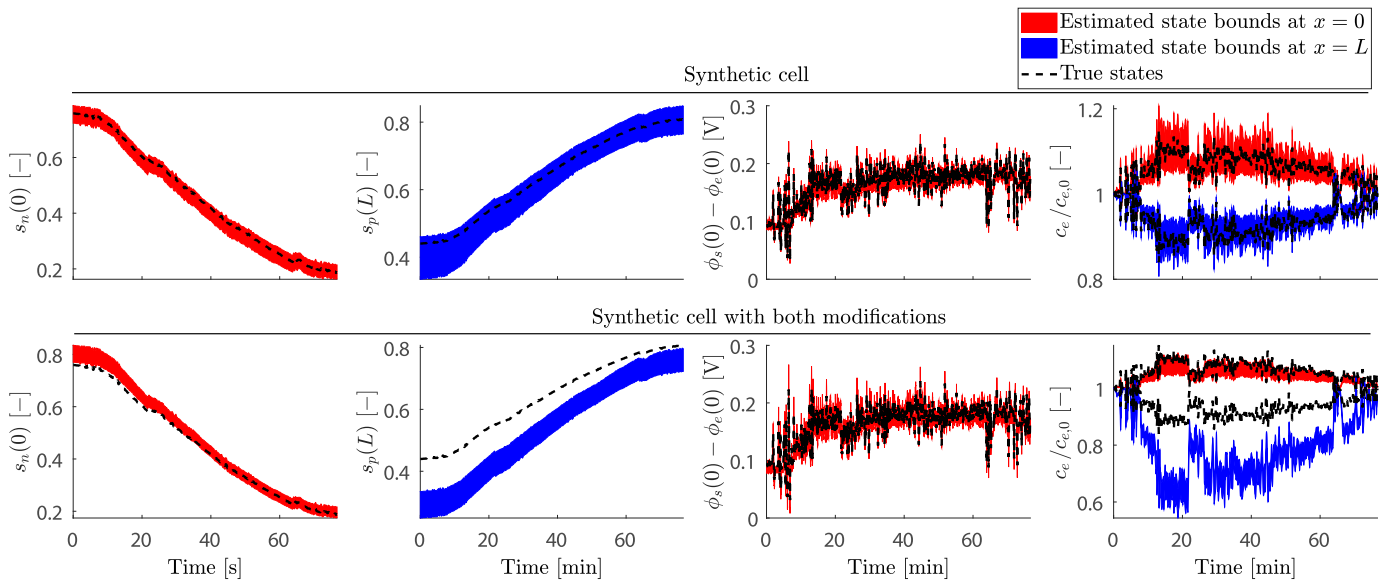

Fig. 8. Upper and lower bounds of the resulting estimated states obtained using the estimated parameter sets.

an inter-dependency between the parameters. Therefore, when the value of a parameter is consistently at the edge of the range, i.e., its estimated values are always close to $\beta = 0$ or $\beta = 1$, as is the case for some parameters in the case of Cell 1, the other parameters will show a smaller variation. We can also observe that especially for the first 10 parameters, the estimated parameters are close to the true parameters, with a relatively small variability. Notably, parameter 5 ($\hat{D}_{s,\text{pos}}$) seems to be the most identifiable parameter, since it has a very small variability and the estimated values are very close to the true value. This implies that the model output is very sensitive to $\hat{D}_{s,\text{pos}}$, while at the same time there is a low inter-dependency of $\hat{D}_{s,\text{pos}}$ with the other parameters.

To see the effect of unmodeled behavior on the estimated parameters, we first add a nonlinearity in $\hat{D}_{s,\text{pos}}$, since this parameter has a large effect on the output. Specifically, we change $\hat{D}_{s,\text{pos}}$ into a concentration-dependent parameter equal to

$$\hat{D}_{s,\text{pos}}(s_p) = 5.705 \times 10^{-4+8(s_p-0.5)^2}.$$

Note that the scaling 5.705 is chosen such that $\hat{D}_{s,\text{pos}}((s_{100,p} + s_{0,p})/2)$ is equal to the originally chosen parameter value for $\hat{D}_{s,\text{pos}}$, since we would like to discount the effect of a bias in the parameter as much as possible. We then also introduce an error in the EMF, which translates into an error in U_p, as

$$U_{\text{EMF}}(s_c) = \bar{U}_{\text{EMF}} + 0.003\sin(4\pi s_c),$$

where $\bar{U}_{\text{EMF}}$ is the original EMF. The various values in the modifications are chosen such that the RMSE between the original model output and the modified model output after parameter estimation is approximately 1.2 mV using the current profile of Data 1, which is in the same range as the errors shown in the upper plot of Fig. 6. The results for these modified models are shown in Fig. 7. We observe that when only adding a nonlinearity in $\hat{D}_{s,\text{pos}}$, apart from parameters 1 and 8, the first 9 parameters are still estimated almost as well as the synthetic cell without any modifications. However, when adding an error in U_p, the estimates deviate significantly more from the true values than when adding a nonlinearity in $\hat{D}_{s,\text{pos}}$, even though the added errors are similar in magnitude.

This is further supported by Table 4, where we can see the median of the RMSE between the estimated parameters and the true parameters. As a reference, the median RMS error (MRMSE) between random sets of parameters and the true parameters is also shown. Here, we observe that the MRMSE of the parameters for the synthetic cell is clearly smaller than the MRMSE of a random set of parameters. This indicates that in the ideal circumstance, when there are no modeling errors, there is a correlation between the closeness of the estimated parameters to the true parameters and the closeness of the resulting output to the true output. However, when adding modeling errors, we can see that this correlation no longer holds. In fact, the MRMSE of the parameters with both of the modifications is almost equal to the MRMSE using a random set of parameters. Interestingly, when computing the MRMSE using only the first 9 parameters, we can see that with a nonlinearity in $\hat{D}_{s,\text{pos}}$, the MRMSE is still significantly lower than the MRSE of random sets of parameters. This means that under this circumstance, with a nonlinearity in $\hat{D}_{s,\text{pos}}$, the first 9 parameters can still be somewhat reliably estimated. With an error in EMF, the MRMSE of the parameters is still close to the MRMSE of random sets of parameters. This suggests that EMF modeling errors affect the identifiability of the parameters significantly more than neglecting the concentration-dependency of parameters. Therefore, when designing experiments for the parameter estimation routine, it is critical that EMF modeling errors are minimized. This might seem obvious as the battery voltage mostly depends on the EMF, see [40], but it demonstrates once more that when designing experiments, it is critical that the EMF is accurately determined.

Table 4
Median of the output RMS error (RMSE) and the median deviation from the true parameters of the obtained estimated parameters from the various synthetic models. The notation $\beta_{\text{est},N}$ indicates that first N parameters have been used for the calculation of the RMSE.

Median of	Synthetic cell	Nonlinearity in $\hat{D}_{s,\text{pos}}$	Error in EMF	Random parameters
RMSE $V_{\text{est}} - V_{\text{true}}$	0.13	1.19	1.17	36.2
RMSE $\beta_{\text{est},22} - \beta_{\text{true}}$	0.30	0.43	0.44	0.43
RMSE $\beta_{\text{est},9} - \beta_{\text{true}}$	0.14	0.22	0.34	0.38

4.5. Consistency and accuracy of the internal states

The results presented above show the consistency and accuracy of the estimated parameters. However, in some applications, such as fast charging of batteries, see e.g., [3–6], only certain internal states are of interest, which are not necessarily significantly affected by all the parameters of the DFN model. Generally, the internal states of interest

are the stoichiometry in the electrodes s_n and s_p, which either are constrained directly, e.g., [5], or are constrained through particle stress models, which depend on the stoichiometry of the electrodes, e.g., [4]. Side-reaction models are also used for the purpose of fast charging, in e.g., [3,6], which mainly depend on the potential drop between the solid and electrolyte phase $\phi_s - \phi_e$ in the negative electrode. Finally, constraints on electrolyte concentration c_e are also used in [5] for the purpose of fast charging.

In Fig. 8, the lower and upper bound of these aforementioned internal states from simulations using the estimated parameters is shown. For the synthetic cell without modifications, we can see that the true states are within the bounds for all internal states. This is no longer the case for the synthetic cell with modifications, where we see that the estimated s_p and c_e in the positive electrode deviate substantially from their respective true states. This also shows that identifiability of the parameters is in principle not a large issue when the desire is to obtain parameters that lead to a good representation of the considered states. Rather, the bigger issue in this case are modeling errors which lead to a biased estimate of the parameters, leading to estimated states that can be far from the true states. This would be an issue when these estimated models are used in fast charging, especially when there are constraints imposed on the electrolyte concentration. In the example of this synthetic cell, the constraint would be activated conservatively, since the estimated positive electrode c_e varies much more in magnitude than the true c_e. This would lead to a conservative use of the battery. On the other hand, it could equally occur that the estimated states show a smaller variation in magnitude, which could lead to the scenario where for the true cell the constraints are violated.

The results shown in this section signify the importance of both the work done in determining the parameters through cell teardown and the work done in determining the parameters based on input/output measurements. As we have shown, using only input/output measurements can lead to a model that does not sufficiently represent the internal states of the cell. On the other hand, we have shown in Fig. 6 that by measuring the parameters, the obtained model does not sufficiently represent the output of the cell. Therefore, in determining a model that both sufficiently represents the internal states and the output, it is crucial that both these approaches in determining model parameters are combined in some way. For example, as we have seen in Fig. 7, determining tighter parameter ranges leads to more consistent parameter estimates, while at the same time limits the deviation that the estimated parameters can have from the true parameters. Note that experimentally determining parameters, as done in e.g., [23] and [24], is thus especially useful in defining tighter parameter ranges, which again shows the usefulness of this approach in combination with parameter estimation based on input/output measurements.

5. Conclusions

In this paper, we have proposed a model parameterization approach of the DFN model, by first reparameterizing the DFN model through normalization and grouping, followed by a sensitivity analysis and a parameter estimation procedure. We have presented results for the parameterization of two cells using experimental data of current and voltage measurements, one where only the EMF (or OCP) measurements of the complete cell have been provided, and one where the individual electrode potentials had been characterized. Here, we have studied the influence of the number of estimated parameters and the length of identification data on the obtained model accuracy. We have found that for both cells, estimating 12 out of all 22 model parameters is sufficient to obtain an accurate model (with respect to the output voltage), and we have found that the length of the identification data should be carefully selected to avoid overfitting of the parameters to modeling errors as much as possible. Finally, we have shown the consistency and accuracy of the parameter estimation routine by analyzing the parameter estimation routine for a synthetic cell. Through this analysis, we have demonstrated that modeling errors, and in particular EMF modeling errors, can lead to a large bias and variability in the estimated parameters. We have further shown that this bias and variability can be reduced by determining tighter parameter ranges, which can be done through cell teardown. The results of this analysis motivate the need for an approach that combines parameter estimation using current/voltage data and parameter estimation through cell teardown. While the former approach can lead to a significantly better model accuracy, it might lead to parameters that are not as physically meaningful as those resulting from the latter approach.

CRediT authorship contribution statement

Z. Khalik: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing - original draft. **M.C.F. Donkers:** Conceptualization, Methodology, Resources, Writing - review & editing, Supervision, Funding acquisition. **J. Sturm:** Investigation, Resources, Writing - review & editing. **H.J. Bergveld:** Conceptualization, Methodology, Resources, Writing - review & editing, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work has received financial support from the Horizon 2020 program of the European Union under the grants ‘Electric Vehicle Enhanced Range, Lifetime And Safety Through INGenious battery management’ (EVERLASTING-713771) and ‘Advancing fail-aware, fail-safe, and fail-operational electronic components, systems, and architectures for fully automated driving to make future mobility safer, affordable, and end-user acceptable’ (AutoDrive-737469).

The authors would like to thank Ishaan Mahashabde, IEC AIR TOOLS PVT LTD, Pune, India for his contribution in carrying out some of the preliminary research that has led to some of the methods and results of this paper.

References

[1] S.J. Moura, F.B. Argomedo, R. Klein, A. Mirtabatabaei, M. Krstic, Battery state estimation for a single particle model with electrolyte dynamics, Trans. Control Syst. Technol. 25 (2) (2017) 453–468, http://dx.doi.org/10.1109/TCST.2016.2571663.

[2] C. Zou, A.G. Kallapur, C. Manzie, D. Nesic, PDE battery model simplification for SOC and SOH estimator design, in: Conf. Decis. & Control, 2015, pp. 1328–1333, http://dx.doi.org/10.1109/CDC.2015.7402395.

[3] R. Klein, N.A. Chaturvedi, J. Christensen, J. Ahmed, R. Findeisen, A. Kojic, Optimal charging strategies in lithium-ion battery, in: Am. Control Conf., 2011, pp. 382–387, http://dx.doi.org/10.1109/ACC.2011.5991497.

[4] B. Suthar, P.W.C. Northrop, R.D. Braatz, V.R. Subramanian, Optimal charging profiles with minimal intercalation-induced stresses for lithium-ion batteries using reformulated pseudo 2-dimensional models, J. Electrochem. Soc. 161 (11) (2014) F3144–F3155, http://dx.doi.org/10.1149/2.0211411jes.

[5] H.E. Perez, X. Hu, S.J. Moura, Optimal charging of batteries via a single particle model with electrolyte and thermal dynamics, in: Am. Control Conf., 2016, pp. 4000–4005, http://dx.doi.org/10.1109/ACC.2016.7525538.

[6] Z. Khalik, H.J. Bergveld, M.C.F. Donkers, Ageing-aware charging of lithium-ion batteries using an electrochemistry-based model with capacity-loss side reactions, in: Am. Control Conf., 2020, pp. 2213–2218, http://dx.doi.org/10.23919/ACC45564.2020.9147404.

[7] B. Yann Liaw, G. Nagasubramanian, R.G. Jungst, D.H. Doughty, Modeling of lithium ion cells—A simple equivalent-circuit model approach, Solid State Ion. 175 (1) (2004) 835–839, http://dx.doi.org/10.1016/j.ssi.2004.09.049.

[8] G.L. Plett, Battery Management Systems, Volume I: Battery Modeling, Artech House Publishers, 2015.

[9] M. Doyle, T.F. Fuller, J. Newman, Modeling of galvanostatic charge and discharge of the lithium/polymer/insertion cell, J. Electrochem. Soc. 140 (6) (1993) 1526–1533, http://dx.doi.org/10.1149/1.2221597.

[10] J. Li, N. Lotfi, R.G. Landers, J. Park, A single particle model for lithium-ion batteries with electrolyte and stress-enhanced diffusion physics, J. Electrochem. Soc. 164 (4) (2017) A874–A883, http://dx.doi.org/10.1149/2.1541704jes.

[11] S. Santhanagopalan, Q. Guo, R.E. White, Parameter estimation and model discrimination for a lithium-ion cell, J. Electrochem. Soc. 154 (3) (2007) A198, http://dx.doi.org/10.1149/1.2422896.

[12] R. Masoudi, T. Uchida, J. McPhee, Parameter estimation of an electrochemistry-based lithium-ion battery model, J. Power Sources 291 (2015) 215–224, http://dx.doi.org/10.1016/j.jpowsour.2015.04.154.

[13] A.M. Bizeray, J.-H. Kim, S.R. Duncan, D.A. Howey, Identifiability and parameter estimation of the single particle lithium-ion battery model, Trans. Control Syst. Technol. 27 (5) (2019) 1862–1877, http://dx.doi.org/10.1109/TCST.2018.2838097.

[14] E. Namor, D. Torregrossa, R. Cherkaoui, M. Paolone, Parameter identification of a lithium-ion cell single-particle model through non-invasive testing, J. Energy Storage 12 (2017) 138–148, http://dx.doi.org/10.1016/j.est.2017.04.008.

[15] H. Pang, L. Mou, L. Guo, F. Zhang, Parameter identification and systematic validation of an enhanced single-particle model with aging degradation physics for Li-ion batteries, Electrochim. Acta 307 (2019) 474–487, http://dx.doi.org/10.1016/j.electacta.2019.03.199.

[16] A.P. Schmidt, M. Bitzer, A.W. Imre, L. Guzzella, Experiment-driven electrochemical modeling and systematic parameterization for a lithium-ion battery cell, J. Power Sources 195 (15) (2010) 5071–5080, http://dx.doi.org/10.1016/j.jpowsour.2010.02.029.

[17] X. Zhou, J. Huang, Impedance-based diagnosis of lithium ion batteries: Identification of physical parameters using multi-output relevance vector regression, J. Energy Storage 31 (2020) 101629, http://dx.doi.org/10.1016/j.est.2020.101629.

[18] N. Jin, D.L. Danilov, P.M. Van den Hof, M. Donkers, Parameter estimation of an electrochemistry-based lithium-ion battery model using a two-step procedure and a parameter sensitivity analysis, Int. J. Energy Res. 42 (7) (2018) 2417–2430, http://dx.doi.org/10.1002/er.4022.

[19] J.C. Forman, S.J. Moura, J.L. Stein, H.K. Fathy, Genetic parameter identification of the Doyle-Fuller-Newman model from experimental cycling of a LiFePO4battery, in: Am. Control Conf., 2011, pp. 362–369, http://dx.doi.org/10.1109/ACC.2011.5991183.

[20] J. Marcicki, M. Canova, A.T. Conlisk, G. Rizzoni, Design and parametrization analysis of a reduced-order electrochemical model of graphite/LiFePO4 cells for SOC/SOH estimation, J. Power Sources 237 (2013) 310–324, http://dx.doi.org/10.1016/j.jpowsour.2012.12.120.

[21] L. Zhang, L. Wang, G. Hinds, C. Lyu, J. Zheng, J. Li, Multi-objective optimization of lithium-ion battery model using genetic algorithm approach, J. Power Sources 270 (2014) 367–378, http://dx.doi.org/10.1016/j.jpowsour.2014.07.110.

[22] Z. Chu, R. Jobman, A. Rodrí guez, G.L. Plett, M.S. Trimboli, X. Feng, M. Ouyang, A control-oriented electrochemical model for lithium-ion battery. Part II: Parameter identification based on reference electrode, J. Energy Storage 27 (2020) 101101, http://dx.doi.org/10.1016/j.est.2019.101101.

[23] M. Ecker, T.K.D. Tran, P. Dechent, S. Käbitz, A. Warnecke, D.U. Sauer, Parameterization of a physico-chemical model of a lithium-ion battery I. Determination of parameters, J. Electrochem. Soc. 162 (9) (2015) A1836–A1848, http://dx.doi.org/10.1149/2.0551509jes.

[24] J. Sturm, A. Rheinfeld, I. Zilberman, F.B. Spingler, S. Kosch, F. Frie, A. Jossen, Modeling and simulation of inhomogeneities in a 18650 nickel-rich, silicon-graphite lithium-ion cell during fast charging, J. Power Sources (ISSN: 0378-7753) 412 (2019) 204–223, http://dx.doi.org/10.1016/j.jpowsour.2018.11.043.

[25] R. Jobman, M.S. Trimboli, G.L. Plett, Identification of lithium-ion physics-based model parameter values, J. Energy Chall. and Mech. 2 (2) (2015) 45–55.

[26] J. Schmalstieg, C. Rahe, M. Ecker, D.U. Sauer, Full cell parameterization of a high-power lithium-ion battery for a physico-chemical model: Part I. Physical and electrochemical parameters, J. Electrochem. Soc. 165 (16) (2018) A3799, http://dx.doi.org/10.1149/2.0321816jes.

[27] V. Ramadesigan, K. Chen, N.A. Burns, V. Boovaragavan, R.D. Braatz, V.R. Subramanian, Parameter estimation and capacity fade analysis of lithium-ion batteries using reformulated models, J. Electrochem. Soc. 158 (9) (2011) A1048, http://dx.doi.org/10.1149/1.3609926.

[28] Z. Khalik, M.C.F. Donkers, H.J. Bergveld, Model simplifications and their impact on computational complexity for an electrochemistry-based battery modeling toolbox, J. Power Sources 488 (2021) 229427, http://dx.doi.org/10.1016/j.jpowsour.2020.229427.

[29] V.R. Subramanian, V. Boovaragavan, V. Ramadesigan, M. Arabandi, Mathematical model reformulation for lithium-ion battery simulations: Galvanostatic boundary conditions, J. Electrochem. Soc. 156 (4) (2009) A260, http://dx.doi.org/10.1149/1.3065083.

[30] B.F. Lund, B.A. Foss, Parameter ranking by orthogonalization—Applied to non-linear mechanistic models, Automatica 44 (1) (2008) 278–281, http://dx.doi.org/10.1016/j.automatica.2007.04.006.

[31] Z. Chu, G.L. Plett, M.S. Trimboli, M. Ouyang, A control-oriented electrochemical model for lithium-ion battery, Part I: Lumped-parameter reduced-order model with constant phase element, J. Energy Storage 25 (2019) 100828, http://dx.doi.org/10.1016/j.est.2019.100828.

[32] L. Xia, E. Najafi, Z. Li, H. Bergveld, M. Donkers, A computationally efficient implementation of a full and reduced-order electrochemistry-based model for Li-ion batteries, Appl. Energy 208 (2017) 1285–1296, http://dx.doi.org/10.1016/j.apenergy.2017.09.025.

[33] K. Smith, C.-Y. Wang, Solid-state diffusion limitations on pulse operation of a lithium ion cell for hybrid electric vehicles, J. Power Sources 161 (1) (2006) 628–639, http://dx.doi.org/10.1016/j.jpowsour.2006.03.050.

[34] H.J. Bergveld, W.S. Kruijt, P.H.L. Notten, Battery Management Systems: Design by Modeling, Kluwer Academic Publishers, 2002.

[35] P. Arora, M. Doyle, A.S. Gozdz, R.E. White, J. Newman, Comparison between computer simulations and experimental data for high-rate discharges of plastic lithium-ion batteries, J. Power Sources 88 (2) (2000) 219–231, http://dx.doi.org/10.1016/S0378-7753(99)00527-3.

[36] M. Doyle, Y. Fuentes, Computer simulations of a lithium-ion polymer battery and implications for higher capacity next-generation battery designs, J. Electrochem. Soc. 150 (6) (2003) A706, http://dx.doi.org/10.1149/1.1569478.

[37] J.N. Reimers, M. Shoesmith, Y.S. Lin, L.O. Valoen, Simulating high current discharges of power optimized Li-ion cells, J. Electrochem. Soc. 160 (10) (2013) A1870–A1884, http://dx.doi.org/10.1149/2.094310jes.

[38] H.P.G.J. Beelen, H.J. Bergveld, M.C.F. Donkers, On experiment design for parameter estimation of equivalent-circuit battery models, in: Conf. Control Technol. and Appl., IEEE, 2018, pp. 1526–1531, http://dx.doi.org/10.1109/CCTA.2018.8511529.

[39] J. Sturm, S. Ludwig, J. Zwirner, C. Ramirez-Garcia, B. Heinrich, M.F. Horsche, A. Jossen, Suitability of physicochemical models for embedded systems regarding a nickel-rich, silicon-graphite lithium-ion battery, J. Power Sources 436 (2019) 226834, http://dx.doi.org/10.1016/j.jpowsour.2019.226834.

[40] J. Sturm, A. Frank, A. Rheinfeld, S.V. Erhard, A. Jossen, Impact of electrode and cell design on fast charging capabilities of cylindrical lithium-ion batteries, J. Electrochem. Soc. 167 (13) (2020) 130505, http://dx.doi.org/10.1149/1945-7111/abb40c.