Applied Energy 321 (2022) 119390

Contents lists available at ScienceDirect

Applied Energy

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

Physics-encoded deep learning in identifying battery parameters without direct knowledge of ground truth

Bin Wu [a,1], Buyi Zhang [a,1], Changyu Deng [a,1], Wei Lu [a,b,*]

[a] Department of Mechanical Engineering, University of Michigan, Ann Arbor, MI 48109, United States
[b] Department of Materials Science and Engineering, University of Michigan, Ann Arbor, MI 48109, United States

HIGHLIGHTS

- Physical laws and on-line observation are embedded into machine learning.
- Low-cost voltage data are used to identify complex battery parameters.
- Machine learning is used without knowledge of ground truth as the training data.
- Electrode diffusivities as complex functions of Li concentration are identified.
- Method is immune to measurement noise and simultaneously estimates many parameters.

ARTICLE INFO

Keywords:
Machine learning
Physics-based model
Parameter estimation
Diffusion
On-line observation

ABSTRACT

We show a method to embed physical laws and on-line observation into machine learning so that irrelevant low-cost battery data can be utilized to identify complex system parameters by machine learning without knowledge of their ground truth as the training data. Lithium diffusivity, a complicated function of lithium concentration, is a crucial parameter for battery performance but difficult to measure directly. We take diffusivity as an example and show that it can be obtained from easily measured sequence of battery voltage over time. In simulations, our results show that this method accurately quantifies not only the diffusivities of both positive and negative electrodes, but also as complex non-linear functions of lithium concentration, purely based on the cell voltage data requiring neither diffusivity nor concentration measurement. Notably, it can accurately predict non-monotonic, many-to-one relations such as "w" shape functions. Moreover, this method is immune to measurement noise and capable of simultaneously estimating multiple parameters. In experiments, our method demonstrates more robust diffusivity estimation than a pure physics-based parameter fitting method and a widely used experimental technique. Our results suggest that the approach enables identifying physical parameters and their interdependence without direct measurements of those parameters.

1. Introduction

The past decades have witnessed great advances in lithium-ion battery technology, which now powers applications from consumer electronics to electric vehicles. Physics based modeling [1–4] has provided comprehensive understanding of the battery behaviors, shedding lights on the battery design optimization. However, the physics-based model consists of dozens of parameters, many of which depend on temperature, state of charge (SOC) or other state variables. Unfortunately, the acquisition of those parameters can be complex and expensive [5–8]. For example, lithium diffusivity, a complicated function of lithium concentration, is a crucial parameter for battery performance [9,10] but difficult to measure directly. The traditional methods to measure lithium diffusivity in active materials, such as electrochemical impedance spectroscopy (EIS) [11], galvanostatic intermittent titration technique (GITT) [12,13] and potential intermittent titration technique (PITT) [14], require expensive electrochemical workstations and sophisticated test recipes. Consequently, the online estimation of physical parameters remains difficult, which limits the applications of physics-based modeling. Meanwhile, modern battery management systems have

* Corresponding author.
E-mail address: weilu@umich.edu (W. Lu).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.apenergy.2022.119390
Received 7 February 2022; Received in revised form 8 May 2022; Accepted 29 May 2022
Available online 14 June 2022

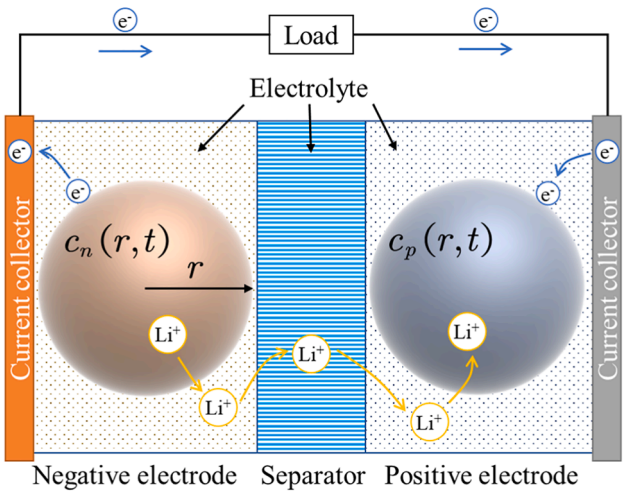

Fig. 1. Illustraion of modeling a discharge step by SPM.

been equipped with multiple voltage, current and temperature sensors and powerful processors, enabling more data to be accessible and analyzed. This requires us to infer crucial parameters from the easily measured voltage, current and temperature signals.

A popular way to model complicated systems is deep learning, which has attracted unprecedented attention from both industry and academia [15–19]. Deep neural networks are well known to be capable of approximating any function [20]. However, pure data-driven models require large amount of data to train [21]. What is worse, even if they are fully trained to predict the relation among signals, the learned models reveal little about the internal mechanisms and are not interpretable [22]. A possible solution may lie in the utilization of existing physics, or more broadly, knowledge [23]. In recent years, researchers have attempted to combine data and physics together for modeling [24]. Physics can be embedded in multiply ways such as features [24] and algorithm structures [25]. A stricter and stronger method is to incorporate partial differential equations (PDE). Zhao *et. al.* [26,27] used PDE-constrained optimization to infer constitutive laws from image data in pattern formation problems. Raissi *et. al.* [28,29] enforced equations to leverage deep neural networks as a PDE solver. For battery state of heath models, Crawford *et. al.* [30] found that incorporating the insights from physics models, their statistic model is faster and more robust for capacity degradation prediction. There are two major issues that were not clearly discussed in the previous studies. One is the validity of the knowledge; they typically took the physics knowledge for granted without investigating the discrepancy between data and physics. The other is the form of unknown parameters; they used constant or predetermined functions to represent unknown physical parameters. We aim to consider the tradeoff between physics and data, and solve for parameters with unknown dependence on other variables. We estimate not only constant parameters, but also the unknown functional relationship between dependent parameters.

In this paper, we introduce a paradigm to infer the internal parameters of a battery system. We present an example of inferring the particle diffusivity as a function of Li ion concentration using only voltage training data, without any knowledge of concentration/diffusivity pair as the training data (in fact, neither concentration nor diffusivity is known). Single particle model (SPM) [31,32] is incorporated into the network as prior knowledge. Our fusion of data-based approach and model-based approach not only helps to reduce the required training data, but also enables the estimation of parameters appearing in physical equations, which can even be the functions of state variables. Thanks to the generality of neural network and our framework, we could estimate any relations between parameters and state variable without any prior knowledge of the function form. For instance, diffusivity is not a constant but depends on the ion concentration, and we can accurately predict the non-monotonic, many-to-one relations such as the "w" shape function of diffusivity as a function of ion concentration, without any diffusivity or concentration data. The prediction accuracy of our method is excellent for simulation data. Furthermore, the performance of our diffusivity estimation is verified using the experimental voltage data. The comparison with the results obtained by two other techniques, direct experimental measurement of diffusivity by GITT and model-based parameter fitting, shows the high stability and reliability of our approach.

2. Methodology

We first introduce a physics-based model, i.e., SPM. Then, we show the proposed approach of incorporating physics-based model and online observation into a data-driven model for model parameter determination by regularizing a deep neural network (DNN) with SPM during training.

2.1. Single particle model

SPM is a simple battery model to describe current and lithium flow in lithium-ion batteries. It has high accuracy when the current is not high. As shown in Fig. 1, the model simplifies the two electrodes as two representative spherical particles and the current is uniform at the particle surfaces. Lithium ion concentration inside the particles is determined by diffusion, which follows the Fick's law,

$$\frac{\partial c_i}{\partial t} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 D_i \frac{\partial c_i}{\partial r}\right) \tag{1}$$

where $c_i(r,t)$ is the solid-phase lithium-ion concentration as a function of the radial coordinate r and time t, subscripts $i = p$ and $i = n$ denote the positive and negative electrode, respectively, and D_i is the diffusion coefficient.

The particle center has zero flux, giving

$$\left.\frac{\partial c_i}{\partial r}\right|_{r=0} = 0 \tag{2}$$

The current density at particle surface ($r = R_i$) is given by

$$-D_i \left.\frac{\partial c_i}{\partial r}\right|_{r=R_i} = J_i \tag{3}$$

where J_i denotes the lithium intercalation/de-intercalation mass flux on the particle surface. The flux is related to the applied current density by

$$J_p = \frac{i_{app}}{Fa_p L_p}, \quad J_n = \frac{-i_{app}}{Fa_n L_n} \tag{4}$$

where i_{app} denotes the applied current density (current per unit area of current collector) and is defined positive during charging (i.e. Li ions flow from the positive electrode to the negative electrode), F denotes the Faraday constant, L_i denotes the electrode thickness, and $a_i = 3\varepsilon_i/R_i$ with ε_i being the volume fraction of the solid phase in the electrode. At $t = 0$, the concentration is known to be

$$c_i(r,0) = c_{i,0} \tag{5}$$

The governing equation (1), boundary conditions (2) and (3), and the initial condition (5) can be used to solve $c_i(r,t)$. However, the battery system is complicated since the diffusivity D_i is very sensitive to internal structure and external environment, and may change as the battery ages. Therefore, it is almost impossible to describe the system *a priori*. We must build an adaptive model with online parameter estimation. In applications, after a battery is assembled, we are not able to directly obtain the lithium concentration, except knowing that the cell voltage, V, depends on the concentration implicitly.

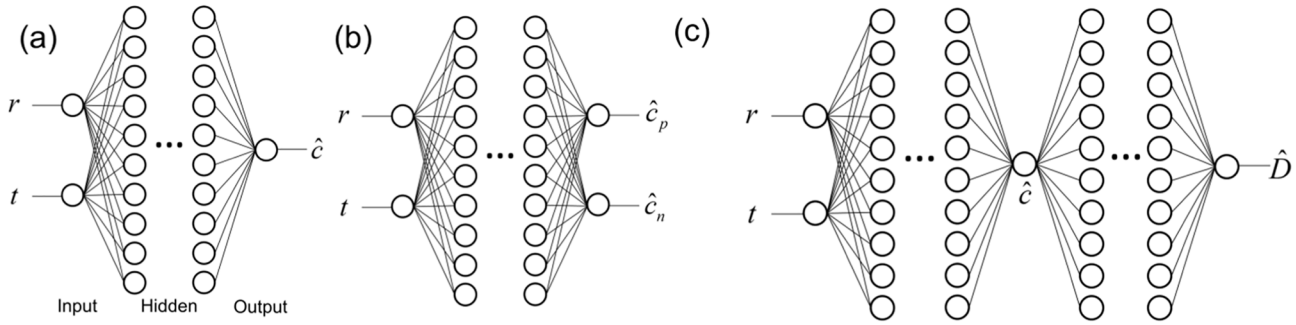

Fig. 2. Schematic of the deep neural networks used for (a) one electrode with an unknown constant diffusivity; (b) two electrodes each having an unknown constant diffusivity; and (c) one or two electrodes with concentration-dependent diffusivity.

Lithium ion transfer current density at the particle surface,FJ_i, is governed by the Butler-Volmer equation,

$$FJ_i = Fk_i\left(c_{i,max} - c_{i,surf}\right)^{0.5} c_{i,surf}^{0.5} c_e^{0.5}\left[\exp\left(\frac{F\eta_i}{2RT}\right) - \exp\left(-\frac{F\eta_i}{2RT}\right)\right] \tag{6}$$

where k_i is the reaction rate constant, $c_{i,max}$ is the maximum lithium concentration in the particle, $c_{i,surf}$ is the lithium ion concentration on the particle surface, c_e is electrolyte concentration, R is ideal gas constant, and T is temperature. $\eta_i = \phi_{s,i} - \phi_{e,i} - U_i(c_{i,surf})$ denotes the overpotential calculated by the solid potential $\phi_{s,i}$, electrolyte potential $\phi_{e,i}$ and the open circuit potential $U_i(c_{i,surf})$ which depends on the surface concentration.

The voltage of the whole cell is given by

$$V = \phi_{s,p} - \phi_{s,n} + i_{app}R_{cell} \tag{7}$$

where R_{cell} is the electrolyte resistance per current collector area. Combining Eqs. (4), (6) and (7), we have

$$V = U_p\left(c_{p,surf}\right) - U_n\left(c_{n,surf}\right) + \frac{2RT}{F}\ln\left(\frac{\sqrt{m_p^2+4}+m_p}{2}\right) + \frac{2RT}{F}\ln\left(\frac{\sqrt{m_n^2+4}+m_n}{2}\right) + i_{app}R_{cell} \tag{8}$$

where $m_p = i_{app}/\left[Fk_pL_pa_p\left(c_{p,max} - c_{p,surf}\right)^{0.5} c_{p,surf}^{0.5} c_e^{0.5}\right]$ and $m_n = i_{app}/\left[Fk_nL_na_n\left(c_{n,max} - c_{n,surf}\right)^{0.5} c_{n,surf}^{0.5} c_e^{0.5}\right]$.

2.2. Embedding physics and on-line observation in neural network

For data-driven models, a straightforward idea to describe the battery system is to train a DNN to learn the function $\hat{c}_i(r,t;\hat{\theta}_i')$ parameterized by learnable weights $\hat{\theta}_i'$ of the network (we use hat to denote learnable parameters or functions). However, the challenge is that we do not have the ground truth of concentration data, and only know the voltage data but Eq. (8) cannot give concentrations backwards.

Now that physics-based SPM and data-driven models are difficult to quantify the battery system individually, we propose to combine them by encoding Eqs. (1)–(5) to a DNN. The diffusivity D_i is regarded as unknown and will be learned together with neuron weights. Thus, we separate $\hat{\theta}_i'$ into $\hat{D}_i$ and $\hat{\theta}_i$, and re-write the learned function approximator as $\hat{c}_i(r,t;\hat{D}_i,\hat{\theta}_i)$, where $\hat{D}_i$ aims to approximate diffusivity D_i and can be a constant or a set of parameters to describe the relation between diffusivity and other variables (such as concentration). By explicitly separating diffusivity from other parameters, we are able to use Eqs. (1)–(5) as a constraint of $\hat{D}_i$. In this way, the approximator is easily trained. Meanwhile, we could extract diffusivity $D_i \approx \hat{D}_i$ after training.

Formally, we train a DNN approximator $\hat{c}_i(r,t;\hat{D}_i,\hat{\theta}_i)$ with a loss function,

$$\begin{aligned} L(\hat{D}_i,\hat{\theta}_i) = {} & w_1 s_{1,i} E_{r,t}\left[\left(\frac{\partial\hat{c}_i}{\partial t} - \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\hat{D}_i\frac{\partial\hat{c}_i}{\partial r}\right)\right)^2\right] + w_2 s_{2,i} E_t\left[\left(\left.\frac{\partial\hat{c}_i}{\partial r}\right|_{r=0}\right)^2\right] \\ & + w_3 s_{3,i} E_t\left[\left(\left.\hat{D}_i\frac{\partial\hat{c}_i}{\partial r}\right|_{r=R_i} + J_i\right)^2\right] + w_4 s_{4,i} E_r\left[\left(\left.\hat{c}_i\right|_{t=0} - c_{i,0}\right)^2\right] \\ & + w_5 s_5 E_t\left[\left(\hat{V} - V\right)^2\right] \end{aligned} \tag{9}$$

where w_1, w_2, …, w_5 are weights and $s_{1,i}$, $s_{2,i}$, …, s_5 are normalization factors to make each term dimensionless, V is the measured voltage response as a function of time, J_i is the preset current density in Eq. (4), $\hat{V}$ is an indirect output of the DNN, i.e., calculated from $\hat{c}_i$ by Eq. (8). $E_t[\cdot]$ denotes temporal average (i.e., $E_t[f(r,t)] = \sum_{i=1}^{N} f(r,t_i)/N$ where N is the number of sampling points in the t axis), $E_r[\cdot]$ denotes spatial average (i.e., $E_r[f(r,t)] = \sum_{i=1}^{N} f(r_i,t)/N$ where N is the number of sampling points in the r axis), and $E_{r,t}[\cdot]$ denotes the average on both space and time (i.e., $E_{r,t}[f(r,t)] = \sum_{i=1}^{N} f(r_i,t_i)/N$ where N is the number of sampling points in the two dimensional r,t plane). For the normalization factors we choose $s_{1,i} = t_{max}^2/\left(c_{i,max} - c_{i,min}\right)^2$ where t_{max} denotes the maximum discharging time considered and $c_{i,min}$ is the minimum lithium concentration in the particle,$s_{2,i} = R_i^2/c_{i,min}^2$,$s_{3,i} = 1/J_i^2$,$s_{4,i} = 1/c_{i,min}^2$, and $s_5 = 1/\left(V_{max} - V_{min}\right)^2$ where V_{max} and V_{min} denote the maximum and minimum measured voltage, respectively.

Note that in Eq. (9) we embed not only the physical law, but also the online observation (measurement data) of V into regularization. The first four terms, which come from the physical equations, would uniquely determine the system if diffusivity were known. Here we assume the material property of diffusivity (as an example of unknown model parameter) is unknown, so that D_i is to be determined. Therefore, we incorporate the online observation data as the fifth term in Eq. (9). This term acts to minimize the difference between predicted and measured voltage data so as to find $\hat{D}_i$. By using the loss function Eq. (9), we not only make it possible to obtain the concentration $\hat{c}_i$, but can also

Table 1
Estimated diffusivity and error with voltage noise levels. The true diffusivity is 1×10^{-14} m^2 s^{-1}. The initial value is set as 1×10^{-13} m^2 s^{-1}, one order of magnitude larger than the true value. Voltage noise is quantified by standard deviation of a Gaussian distribution.

Voltage noise (V)	Estimated diffusivity (m^2 s^{-1})	Error
0	1.00089×10^{-14}	0.08901%
0.005	1.00534×10^{-14}	0.53378%
0.01	9.80229×10^{-15}	1.97707%
0.03	9.75887×10^{-15}	2.41130%

Table 2
Estimated diffusivity of positive and negative electrodes and electrolyte resistance. The initial values for D_p and D_n are 1×10^{-13} m^2 s^{-1}, and for R_{cell} is 1×10^{-3} Ω m^2.

Parameter	Truth	Estimated	Error
D_p	1×10^{-14} m^2 s^{-1}	1.00412×10^{-14} m^2 s^{-1}	0.41240%
D_n	1×10^{-14} m^2 s^{-1}	1.00317×10^{-14} m^2 s^{-1}	0.31719%
R_{cell}	3.24×10^{-4} Ω m^2	3.23483×10^{-4} Ω m^2	0.15953%

extract the diffusivity $\widehat{D}_i$ as a function of concentration without requiring knowing any prior knowledge of their relation. The dependence of $\widehat{D}_i$ on $\widehat{c}_i$ can be any form and we use a neural network to describe the relation. Examples and details of DNN are given in the following section.

3. Simulation

We use four simulation examples to demonstrate our method and test its performance. Two of the examples use a half cell where the positive electrode is porous lithium manganese oxide (LMO) and the negative electrode is lithium metal. Since the reaction at the lithium metal has an equilibrium potential of 0 V with respect to Li^+/Li and the voltage loss of this reaction is assumed negligible comparing to the positive electrode, we can set $U_n = 0$ and $m_n = 0$ in Eq. (8). In the other two examples, we simulate a full cell with LMO as the positive electrode and graphite as the negative electrode. In our simulation, we discharge a fully charged battery until a certain voltage threshold. Typical battery properties and discharge conditions can be found in ref [5]. We use the finite element software package, COMSOL Multiphysics, to simulate the voltage of a cell with various given diffusivity (as functions of concentrations) during discharging. The simulated voltage as a function of time, treated as the measured signal, serves as the training data of a DNN. The diffusivity functions used in generating the simulated voltage data are the ground truth diffusivity, which will be used to access the accuracy of the estimated diffusivity.

DNN is in general used to approximate $\widehat{c}_i(r, t; \widehat{D}_i, \widehat{\theta}_i)$, as shown in Fig. 2. There are some variances of the inputs and outputs, which are detailed later in the examples. In Eq. (9), by which the DNN is trained, we put different weights on different terms. A larger weight is assigned to the last term of Eq. (9) by $w_5 = 10$. Other weights are $w_{1,i} = 2$ and $w_{2,i} = w_{3,i} = w_{4,i} = 1$. In each training step, 4,000 spatial–temporal points are sampled from the space $[0, R_i] \times [0, t_{max}]$ to calculate the expectation in the loss function. The loss function is minimized by the L-BFGS-B algorithm for 50,000 steps.

3.1. One electrode with unknown constant diffusivity

We start with a half cell with a lithium metal reference/counter electrode and a positive porous electrode whose diffusivity is constant. A deep neural network $\widehat{c}(r, t; \widehat{D}, \widehat{\theta})$ is used to represent the relationship between the inputs r (distance from particle center), t (time) and output $\widehat{c}$ (lithium concentration in the positive electrode), as shown in Fig. 2a. This is a half cell so for simplicity we omit the subscript 'p'. The DNN is composed of eight intermediate linear layers which are all 20-dimensional and connected by the tanh activation function; positive electrode diffusivity $\widehat{D}$ is learned together with DNN weights $\widehat{\theta}$. The essence to achieve a better approximation of $\widehat{c}$ to the true concentration, c, lies in the carefully designed loss function $L(\widehat{D}, \widehat{\theta})$ in Eq. (9).

To show the effectiveness and robustness of the proposed method, we add Gaussian noise to the simulated voltage. As listed in Table 1, the estimation results show that the developed method is accurate despite the presence of measurement noise.

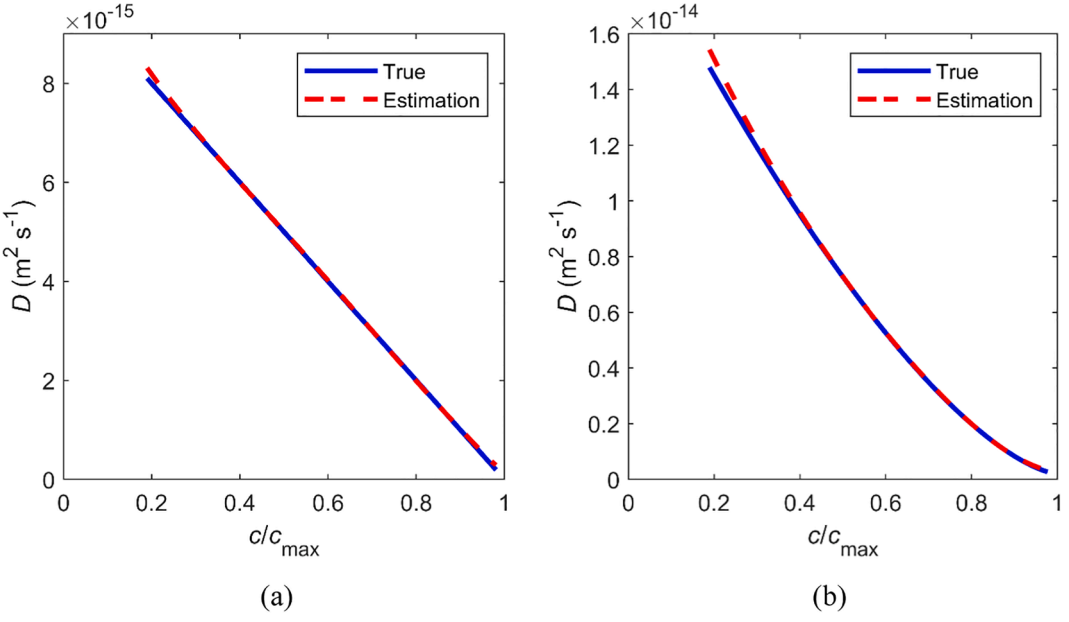

Fig. 3. Estimation results of concentration-dependent diffusivity for a half-cell. The ground truth diffusivity of positive electrode is (a) $D(c) = D_0(1 - c/c_{max})$, where $D_0 = 10^{-14}$ m^2 s^{-1} (b) $D(c) = D_0\left(1 + 100 \times (1 - c/c_{max})^{1.5}\right)$, where $D_0 = 2 \times 10^{-16}$ m^2 s^{-1}. The form of $D(c)$ is unknown to the DNN.

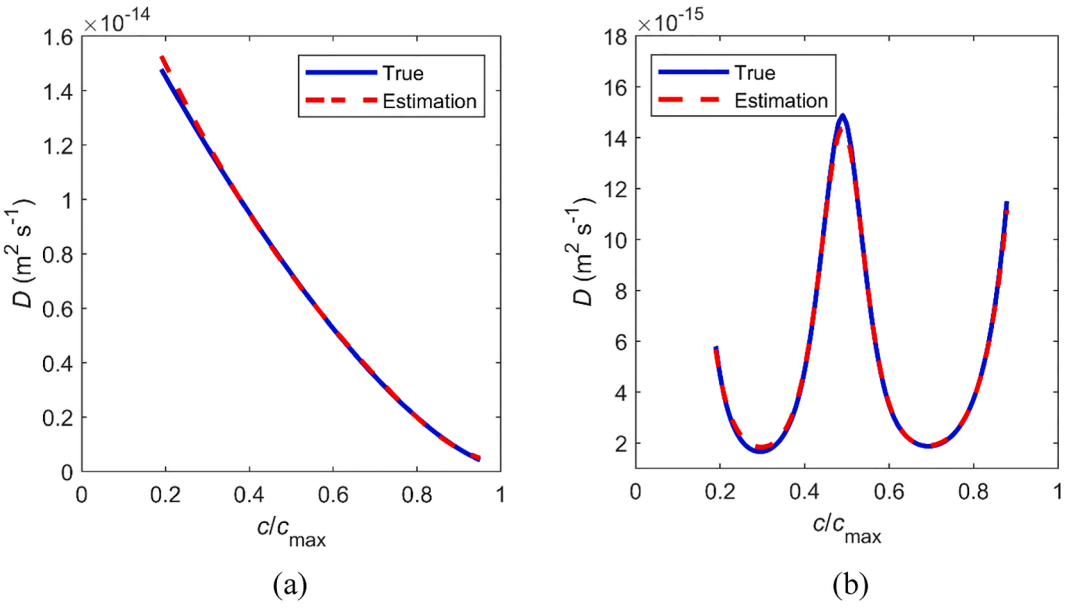

Fig. 4. Estimated concentration-dependent diffusivity for the positive electrode. The negative electrode diffusivity is set as $D(c) = D_0(1 - c/c_{max})$ with D_0 unknown. The ground truth positive electrode diffusivity is (a) a power law function, and (b) a w-shaped function. The form of $D(c)$ is unknown to the DNN.

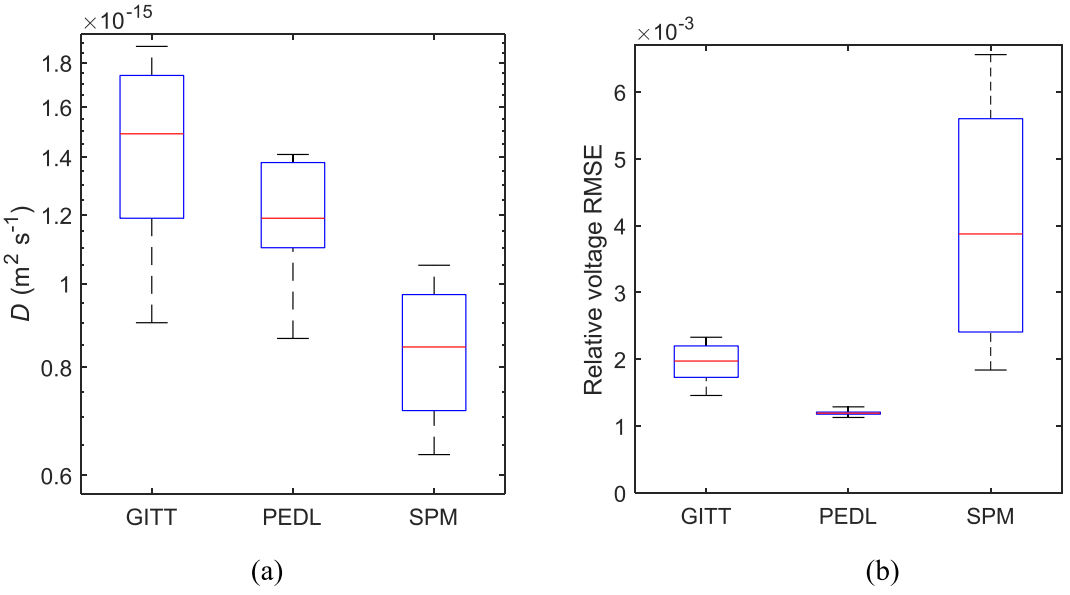

Fig. 5. Box plots of (a) diffusivity estimation results and (b) relative voltage root mean squared error (RMSE) of galvanostatic intermittent titration technique (GITT), our proposed physics-encoded deep learning (PEDL) and single particle model (SPM) with voltage data in different time lengths. The center lines denote median values, the boxes denote upper and lower quantiles, and the two outer lines denote extreme values.

3.2. Two electrodes with unknown constant diffusivity and electrolyte resistance

We extend the method to estimate multiple parameters. Using the neural network in Fig. 2b, we applied the method to not only estimating the diffusivity of both positive and negative electrodes in a full cell, but also the electrolyte resistance. Two DNNs, $\widehat{c}_n(r,t;\widehat{D}_n,\widehat{R}_{cell},\widehat{\theta}_n)$ and $\widehat{c}_p(r,t;\widehat{D}_p,\widehat{R}_{cell},\widehat{\theta}_p)$ are both composed of eight intermediate linear layers which are all 20-dimensional and are connected by tanh activation function. The results listed in Table 2 show that the developed method is capable of simultaneously estimating multiple parameters with good accuracy.

3.3. One electrode with unknown concentration-dependent diffusivity

Assuming that lithium diffusivity is a function of concentration, we introduce an additional sequential deep neural network as shown in Fig. 2c in order to capture the relation between concentration and diffusivity.

There are two DNNs in a sequence, $\widehat{c}_p(r,t;\widehat{\theta}_p)$ and $\widehat{D}_p(c_p;\widehat{\theta}_p')$. The first one is composed of eight intermediate linear layers which are all 20-dimensional and connected by the tanh activation function, and the second one is composed of three 5-dimensional linear layers connected by the tanh activation function. This sequential layout of neural networks ensures the diffusivity to be only dependent on concentration. Note that in constructing the second neural network, we have no pre-

existing knowledge on the explicit form of function relating diffusivity to concentration. The layout simultaneously solves parameters to minimize the loss function, which satisfies the governing laws, while enforcing diffusivity to be a function of concentration. The function is automatically learned through minimizing the loss defined in Eq. (9). Fig. 3 shows the estimated diffusivity in two examples with two different forms. Both Fig. 3a and Fig. 3b show good agreement between the estimated and true values, regardless of diffusivity's linear or non-linear dependence on concentration.

3.4. Two electrodes with unknown concentration-dependent diffusivity

An even more challenging task is to estimate the diffusivities of both electrodes as functions of concentration at the same time in a full cell. In this case, we assume that the positive electrode diffusivity has an unknown relation with concentration, while the negative electrode diffusivity follows a linear relation with concentration, i.e.$D(c) = D_0(1 - c/c_{max})$, but D_0 is unknown. We use three DNNs,$\widehat{c}_n(r, t; \widehat{\theta}_n, \widehat{D}_0)$, $\widehat{c}_p(r, t; \widehat{\theta}_p)$ and $\widehat{D}_p(c_p; \widehat{\theta}_p')$. The first two have the same architecture, eight intermediate linear layers which are all 15-dimensional and connected by the tanh activation function, and the last one is composed of three 5-dimensional linear layers connected by tanh activation function. Negative electrode diffusivity D_0 is a parameter learned together with DNN weights.

We use two examples to highlight the power of this method. In the first example, the positive electrode diffusivity is set as a power law function. The estimated positive electrode diffusivity result is shown in Fig. 4a. The unknown coefficient D_0 in the negative electrode diffusivity relation is estimated as 9.9817×10^{-15} m^2 s^{-1} while the true value is 1×10^{-14} m^2 s^{-1}. In the second example, the positive electrode diffusivity is set as a complicated non-monotonic, many-to-one relation "w" shape function of concentration. The estimation result is shown in Fig. 4b. The unknown coefficient D_0 in the negative electrode diffusivity relation is estimated as 9.9791×10^{-15} m^2 s^{-1} while the true value is 1×10^{-14} m^2 s^{-1}. The training loss and concentration profiles are shown in Appendix A. Both examples indicate good agreement between the DNN-extracted results and the true values.

4. Experiment

To further elucidate the performance of our proposed approach, we estimate the solid diffusivity of graphite in a lithium-graphite coin cell. We also compare the results of our method with direct experimental measurement by GITT and SPM-based parameter fitting.

4.1. Cell fabrication

We made CR2032 coin cells with lithium and graphite electrodes. Graphite powder was mixed with polyvinylidene fluoride (PVDF) binder (10 wt%), super P (10 wt%) and 1-Methyl-2-pyrrolidone to make a homogeneous slurry. The slurry was pasted on Cu foil and vacuum-dried at 110 °C for 12 h. Next, graphite electrode pieces were cut into discs with a diameter of 10 mm. The mass loading was about 4 mg cm^{-2}. The graphite disc, lithium metal counter electrode and separator were sealed and assembled into 2032 type coin cells in an argon-filled glove box with less than 0.1 ppm oxygen and moisture. The electrolyte solution was 1 M lithium hexafluorophosphate solution ($LiPF_6$) dissolved in a mixture (1:1, v/v) of ethylene carbonate (EC) and dimethyl carbonate (DMC).

4.2. Measurement

We firstly cycled the assembled cell 10 times at 50 mA $g^{-1}_{graphite}$ between 0.01 V and 1.5 V to stabilize the performance. Then the cell was rested for at least 12 h to reach equilibrium. To perform graphite diffusivity estimation, we applied a pulse charging (graphite deintercalation) current of 50 mA $g^{-1}_{graphite}$ to the cell with a computer-controlled MACCOR cycler. The device collected the voltage data of the cell for 30 s.

4.3. Postprocessing and results

In addition to our method, we use GITT and SPM to obtain the diffusivity value for comparison.

GITT is a standard approach to measure diffusivity by applying a current pulse, and the measured diffusivity is given by [12]

$$D = \frac{4}{\pi}\left(\frac{IV_M}{SF}\right)^2\left[\left(\frac{\mathrm{d}U^0}{\mathrm{d}\delta}\right)\Big/\left(\frac{\mathrm{d}V}{\mathrm{d}\sqrt{t}}\right)\right]^2 \tag{10}$$

where V_M is the molar volume of the sample, V is the cell voltage as a function of time t after the current pulse, $dU^0/d\delta$ denotes the derivative of the open circuit voltage with respect to state of charge (SOC) when the full stoichiometry window is used, I is the total current, S is the total surface area and F is the Faraday constant.

For SPM-based parameter fitting, we built a single particle model using COMSOL which solves Eqs. (1)–(8). The diffusivity D is estimated from the charging curve fitting using the optimization module. R_{cell} is pre-determined from the first voltage point and cell open circuit voltage using Eq. (8).

For our proposed method, the problem is simplified as one electrode with unknown constant diffusivity. We reduced the loss function weight w_5 to 2 and keep other weights the same ($w_{1,i} = 2$, $w_{2,i} = w_{3,i} = w_{4,i} = 1$). R_{cell} is pre-determined from the first voltage point and cell open circuit voltage using Eq. (8).

The material properties used in the models are presented in Appendix B.

We used voltage data in six different time lengths, 5 s, 10 s, 15 s, 20 s, 25 s and 30 s, and they all starts from 0 s. For GITT, we removed the first three data points to eliminate capacitive effects [12]; for other methods, we used all data points in the interval. As expected, the fitted diffusivity value varies with time lengths. The distributions of estimated diffusivity from three different methods are shown in Fig. 5a. Taking the GITT results as a reference, our method (denoted as PEDL for physics encoded deep learning) shows better stability towards data selection: the estimated diffusivity has the smallest variance between using the six different time lengths. This is because (1) GITT is only accurate in a short time period [12] and the SPM encoded in neural network is valid in a broader time range than Eq. (10) used in the GITT technique, and (2) our method balances the error from equations and data so it is more immune to violation of physical assumptions.

Meanwhile, our method is more accurate than the traditional SPM solved by finite element method (FEM). The cause is similar to the second reason for GITT. In SPM, the initial condition is that Li concentration is assumed to be uniformly distributed across the radius. Yet, that may not be true in experiment. Also, due to variance of particle shape and radius, the voltage prediction of SPM is not accurate. Our method shows the superior capability to correct the inaccuracy in model assumptions by leveraging the voltage data. It estimates the diffusivity by balancing data-driven and physics-based models. When the model assumptions are not accurate and lead to large data discrepancy, our method could relax the assumptions and get more satisfying estimation results. In contrast, the parameter estimation based on physics-based models suffer from the uncertainty of model assumptions such as the initial conditions and constant cycling current because those conditions need to be strictly satisfied for the models to be work. To compare the voltage difference between model predictions and experimental data, we calculated the relative root mean squared error (RMSE) defined by

$$\text{relative RMSE} = \sqrt{\sum_{i=1}^{N}\frac{(V_i - \widehat{V}_i)^2}{NV_i^2}} \tag{11}$$

The results are shown in Fig. 5b. The SPM-based parameter fitting performs the worst with a high error. Detailed data are available in Appendix C.

5. Conclusions

We have developed a systematic approach to estimate battery parameters (diffusivity as an example) through physics-encoded deep learning trained with easily measured voltage data. We successfully estimated the electrode diffusivity as a general function of concentration, which was experimentally unobtainable. Moreover, the accuracy and stability of our method are verified by the experimental voltage data. The proposed framework that integrates machine learning, physics-based modeling, and on-line observation provides a powerful tool capable of identifying the values of parameters and their interdependence functions in complex physical systems.

Author contributions

B.W. and B.Z. built the model, carried out simulations and performed analysis. C.D. conducted the experiment and contributed to validation and analysis of results. W. L. conceptualized and supervised this study. All authors contributed to the writing and editing of the manuscript.

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.

Acknowledgements

The authors gratefully acknowledge the support by LG Energy Solution.

Appendix A. Additional details of the example with w-shape diffusivity in Section 3.4

Some additional details about the example with w-shape diffusivity (Fig. 4b) are shown below. Fig. A1 shows the training loss, which

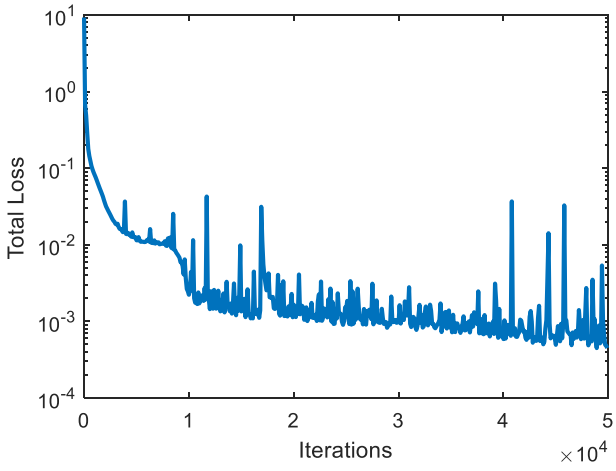

Fig. A1. Evolution of training loss.

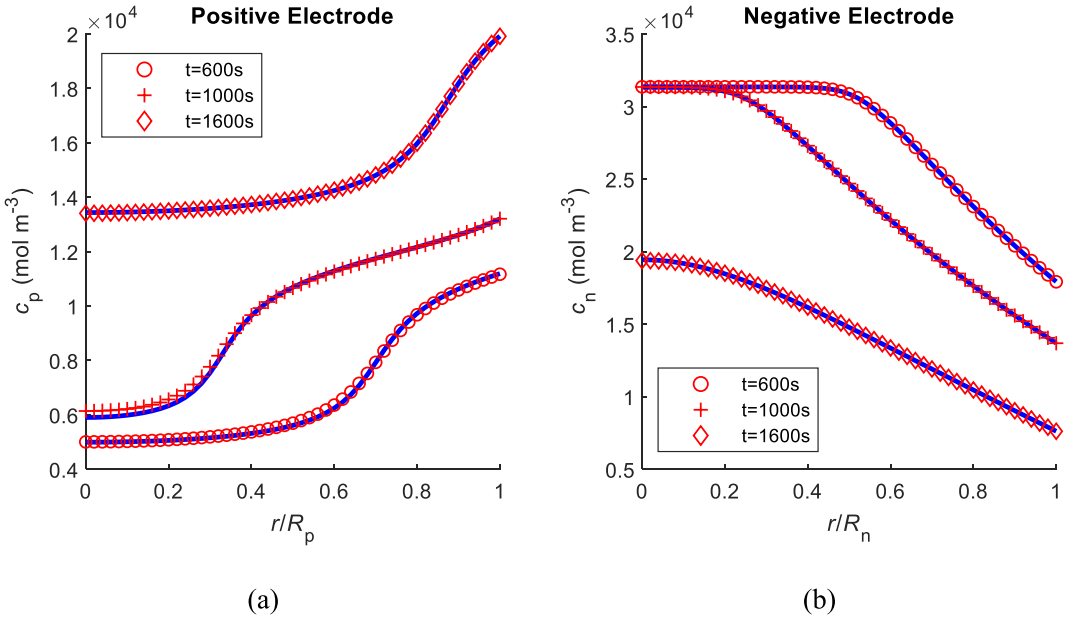

Fig. A2. Concentration profiles of (a) positive electrode and (b) negative electrode. The marks denote the predictions by DNN. The solid blue lines denote the ground truth calculated by SPM. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

confirms good convergence. The concentration profiles of the two electrodes are shown in Fig. A2. The predicted concentration distributions at different time steps (marks) agree well with the ground truth (solid blue lines). We can observe that the concentration of the positive electrode is a complicated function due to the complexity of w-shape diffusivity.

Appendix B. Parameters used to fit experiment data

The material properties used in Section 4 are presented in Tables B1 and B2.

Table B1
Parameters used in SPM and PEDL.

Symbol	Description	Value
c_{max}	Maximum solid Li concentration	31,507 mol m^{-3}
r_p	Graphite particle radius	10.8 μm
ϵ_s	Solid volume fraction	0.28
c_e	Electrolyte concentration	1000 mol m^{-3}
k	Reaction rate constant	5×10^{-10} m$^{2.5}$ mol$^{-0.5}$ s^{-1}
T	Temperature	298 K
l	Graphite electrode thickness	30 μm
I	Applied current	1.2×10^{-4} A
A_{cell}	Cell/electrode cross-sectional area	7.854×10^{-5} m

Table B2
Parameters used in GITT.

Symbol	Description	Value
V_M	Molar volume of lithium in graphite	3.17×10^{-5} m^3 mol^{-1}
S	Graphite surface area,$A_{cell} l \times 3\epsilon_s / r_p$	1.83×10^{-4} m^2
I	Applied current	1.2×10^{-4} A

Appendix C. Fitted diffusivity values and relative error on experimental data

The fitted diffusivity values and relative error are presented in Tables C1 and C2 respectively.

Table C1
Estimated diffusivity D from different methods and data selection (unit: m^2 s^{-1}).

Data time length (s)	GITT	PEDL	SPM
5	9.01E-16	1.15E-15	6.34E-16
10	1.19E-15	8.64E-16	7.13E-16
15	1.40E-15	1.10E-15	8.02E-16
20	1.58E-15	1.23E-15	8.87E-16
25	1.74E-15	1.41E-15	9.71E-16
30	1.88E-15	1.38E-15	1.05E-15

Table C2
Relative RMSE from different methods and data selection.

Data time length (s)	GITT	PEDL	SPM
5	1.46E-03	1.13E-03	1.84E-03
10	1.73E-03	1.29E-03	2.41E-03
15	1.87E-03	1.21E-03	3.25E-03
20	2.08E-03	1.18E-03	4.50E-03
25	2.20E-03	1.19E-03	5.60E-03
30	2.33E-03	1.20E-03	6.56E-03

References

[1] Doyle M, Fuller TF, Newman J. Modeling of galvanostatic charge and discharge of the lithium polymer insertion cell. J Electrochem Soc 1993;140:1526–33.
[2] Tanim TR, Rahn CD. Aging formula for lithium ion batteries with solid electrolyte interphase layer growth. J Power Sources 2015;294:239–47.
[3] Lin X, Park J, Liu L, Lee Y, Sastry AM, Lu W. A comprehensive capacity fade model and analysis for li-ion batteries. J Electrochem Soc 2013;160(10):A1701–10.
[4] Yang XG, Ge SH, Liu T, Leng YJ, Wang CY. A look into the voltage plateau signal for detection and quantification of lithium plating in lithium-ion cells. J Power Sources 2018;395:251–61.
[5] Wu B, Lu W. A consistently coupled multiscale mechanical–electrochemical battery model with particle interaction and its validation. J Mech Phys Solids 2019;125:89–111.
[6] Ecker M, Käbitz S, Laresgoiti I, Sauer DU. Parameterization of a physico-chemical model of a lithium-ion battery II. model validation. J Electrochem Soc 2015;162(9):A1849–57.
[7] Lin X, Lu W. A framework for optimization on battery cycle life. J Electrochem Soc 2018;165(14):A3380–8.
[8] Berliner MD, Zhao HB, Das S, Forsuelo M, Jiang BB, Chueh WH, Bazant MZ, Braatz RD. Nonlinear identifiability analysis of the porous electrode theory model of lithium-ion batteries. J Electrochem Soc. 2021;168.
[9] Goodenough JB. How we made the Li-ion rechargeable battery. Nat Electron 2018;1:204.
[10] Tian RY, Park SN, King PJ, Cunningham G, Coelho J, Nicolosi V, et al. Quantifying the factors limiting rate performance in battery electrodes. Nat Commun 2019;10:1933.
[11] Takami N, Satoh A, Hara M, Ohsaki I. Structural and kinetic characterization of lithium intercalation into carbon anodes for secondary lithium batteries. J Electrochem Soc 1995;142:371–9.
[12] Deng C, Lu W. Consistent diffusivity measurement between Galvanostatic Intermittent Titration Technique and Electrochemical Impedance Spectroscopy. J Power Sources 2020;473:228613.
[13] Yu P, Popov BN, Ritter JA, White RE. Determination of the lithium ion diffusion coefficient in graphite. J Electrochem Soc 1999;146:8–14.
[14] Guyomard D, Tarascon JM. Li metal-free rechargeable $LiMn_2O_4$/carbon cells - their understanding and optimization. J Electrochem Soc 1992;139:937–48.
[15] Carrasquilla J, Melko RG. Machine learning phases of matter. Nat Phys 2017;13:431.
[16] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436–44.
[17] Falk T, Mai D, Bensch R, Cicek O, Abdulkadir A, Marrakchi Y, et al. U-Net: deep learning for cell counting, detection, and morphometry. Nat Methods 2019;16:67–70.
[18] Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, et al. A guide to deep learning in healthcare. Nat Med 2019;25:24–9.
[19] Reichstein M, Camps-Valls G, Stevens B, Jung M, Denzler J, Carvalhais N. Prabhat, Deep learning and process understanding for data-driven Earth system science. Nature 2019;566:195–204.
[20] Goodfellow I, Bengio Y, Courville A. Deep Learning. MIT Press; 2016.
[21] Al-Jarrah OY, Yoo PD, Muhaidat S, Karagiannidis GK, Taha K. Efficient machine learning for big data: A review. Big Data Res 2015;2(3):87–93.
[22] Dawson-Elli N, Lee SB, Pathak M, Mitra K, Subramanian VR. Data science approaches for electrochemical engineers: an introduction through surrogate model development for lithium-ion batteries. J Electrochem Soc 2018;165(2):A1–15.
[23] Deng C, Ji X, Rainey C, Zhang J, Lu W. Integrating machine learning with human knowledge. iScience 2020;23:101656.
[24] Mistry A, Franco AA, Cooper SJ, Roberts SA, Viswanathan V. How machine learning will revolutionize electrochemical sciences. ACS Energy Lett 2021;6:1422–31.
[25] Finegan DP, Zhu J, Feng X, Keyser M, Ulmefors M, Li W, et al. The application of data-driven methods and physics-based learning for improving battery safety. Joule 2021;5(2):316–29.
[26] Zhao H, Braatz RD, Bazant MZ. Image inversion and uncertainty quantification for constitutive laws of pattern formation. J Comput Phys 2021;436:110279.
[27] Zhao HB, Storey BD, Braatz RD, Bazant MZ. Learning the physics of pattern formation from images. Phys Rev Lett 2020;124:060201.
[28] Raissi M, Perdikaris P, Karniadakis GE. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J Comput Phys 2019;378:686–707.
[29] Raissi M, Wang Z, Triantafyllou MS, Karniadakis GE. Deep learning of vortex-induced vibrations. J Fluid Mech 2019;861:119–37.
[30] Crawford AJ, Choi D, Balducci PJ, Subramanian VR, Viswanathan VV. Lithium-ion battery physics and statistics-based state of health model. J Power Sources 2021;501:230032.
[31] Santhanagopalan S, Guo QZ, Ramadass P, White RE. Review of models for predicting the cycling performance of lithium ion batteries. J Power Sources 2006;156:620–8.
[32] Guo M, Sikha G, White RE. Single-particle model for a lithium-ion cell: thermal behavior. J Electrochem Soc 2011;158(2):A122.