



Machine learning-based fast charging of lithium-ion battery by perceiving and regulating internal microscopic states

Zhongbao Wei^{a,*}, Xiaofeng Yang^a, Yang Li^b, Hongwen He^a, Weihai Li^c, Dirk Uwe Sauer^c

^a National Engineering Research Center of Electric Vehicles, School of Mechanical Engineering, Beijing Institute of Technology, Beijing 100081, China

^b Department of Electrical Engineering, Chalmers University of Technology, 41296 Gothenburg, Sweden

^c Chair for Electrochemical Energy Conversion and Storage Systems, Institute for Power Electronics and Electrical Drives (ISEA), RWTH Aachen University, Jaegerstrasse 17/19, 52066, Aachen, Germany

ARTICLE INFO

Keywords:

Fast charging
Machine learning
Battery-to-cloud
Soft actor-critic
Lithium-ion battery

ABSTRACT

Fast charging of the lithium-ion battery (LIB) is an enabling technology for the popularity of electric vehicles. However, high-rate charging regardless of the physical limits can induce irreversible degradation or even hazardous safety issues to the LIB system. Motivated by this, this paper proposes a machine learning-based fast charging strategy with multi-physical awareness within a battery-to-cloud framework. In particular, a reduced-order electrochemical-thermal model is built in the cloud to perceive the microscopic states of LIB, leveraging which the soft actor-critic (SAC) deep reinforcement learning (DRL) algorithm is exploited for the first time to train a fast charging strategy. Hardware-in-Loop tests and experiments with practical LIBs are carried out for validation. Results suggest that the battery-to-cloud architecture can mitigate the risk of a heavy computing burden in the real-time controller. The proposed strategy can effectively mitigate the unfavorable over-temperature and lithium deposition, which benefits the safety and longevity during fast charging. Given a similar charging speed, the proposed machine learning approach extends the LIB cycle life by about 75% compared to the commonly-used empirical protocol. Meanwhile, the proposed strategy is proven superior to the state-of-the-art rule-based and the model-based strategies in terms of charging rapidity, charging safety and computational complexity. Moreover, the trained low-complexity strategy is highly adaptive to the ambient temperature and initial charging state, which promises robust performance in practical applications.

1. Introduction

Lithium-ion batteries (LIBs) have seen wide applications in electric vehicles (EVs) attributed to their advantageous properties of long service life, high gravimetric and volumetric densities [1,2]. Regarding LIB utilization, fast charging is recognized as an enabling technique unlocking the obstacle of slow refueling of EVs compared with the gasoline-powered vehicles [3,4]. However, the partial pursuit of high-rate charging potentially violates the physical constraints of LIB, and thus risks triggering several unfavorable side reactions. This can further incur the quick irreversible degradation and even safety hazards of LIBs [5]. Therefore, the LIB charging is a challenging task that should meet the conflicting objectives of charging rapidity and safety.

The heuristic rule-based methodologies are widely adopted in real-world applications. Prominent delegates include the constant-current-constant-voltage (CCCV) charging protocol [6] and its variants, such

as the multistage constant current-constant voltage (MCC—CV) [7–9], CV-CCCV [10] and pulse charging (PC) [11,12]. These approaches can be easily deployed by predefining charging profiles with fixed current or voltage constraints. However, the inner multi-physical dynamics of the battery have been mostly ignored. Hence, such protocols lack a theoretical guarantee for the optimality with respect to the charging rapidity, safety and longevity.

By comparison, the model-based charging strategies manifest themselves with the possibility for describing and constraining the immeasurable inner states of LIB. Leveraging a coupled electro-thermal (CET) model, a multi-objective evolutionary approach was employed to determine the optimal charging pattern balancing the charging time, energy loss and the temperature build-up [13]. Based on an electro-thermal-aging model, an optimized MCC charging strategy was proposed to constrain the aging stresses and thus elongate the battery life [14]. Attributed to the refined interpretation of the multi-physical

* Corresponding author.

E-mail address: weizb@bit.edu.cn (Z. Wei).

<https://doi.org/10.1016/j.ensm.2022.12.034>

Received 6 May 2022; Received in revised form 6 December 2022; Accepted 20 December 2022

Available online 7 January 2023

2405-8297/© 2022 Elsevier B.V. All rights reserved.

dynamics of LIB, the electrochemical model also appeals for the model-based charging optimization. In [15], a health-aware charging method was developed using a pseudo-2-D electrochemical model, where the capacity fade, SEI growth and over-temperature were suppressed by controlling the relevant inner states within the expected thresholds. Depending on an EM with side reaction models, an optimal charging protocol was designed to trade-off the degradation rate and the charging time of LIB [16]. The charging trajectories are generated before practical implementations via offline optimization in afore-mentioned methods, and thus, they can also be called as trajectory generators. In spite of the enhanced optimality over the conventional rule-based strategies, the trajectory generator as an offline approach also lacks the robustness to the change of operating conditions.

In contrast to the trajectory generators, model-based online controllers give optimized charging decisions in real time along with the acquisition of onboard battery parameters. Relying on the ECM and lumped thermal model, the model predictive control (MPC) have been employed to regulate the electrical and thermal states of LIB within the safety thresholds during the fast charging [17,18]. To probe into and regulate the inner microscopic states of LIB, a proportional-integral-derivative controller was explored to control the anode potential and decline the lithium deposition, with the aid of a simplified pseudo-2-D model [19]. The ion concentrations of LIB electrolytes were further incorporated and constrained via the MPC to prohibit the detrimental side reactions [20]. A nonlinear programming-based charging strategy was developed in [21], where the electrolyte concentration, anode over-potential and internal temperature were regulated simultaneously to ensure the safety- and health-aware charging. Attributed to the closed-loop feedback mechanism, the model-based online charging strategies promise stronger robustness against the external disturbances.

However, the model-based optimization is computationally heavy due to the involvement of high-dimensional matrix manipulation and intractable nonlinear optimization. It is also foreseeable that the computing burden will further build up exponentially with the use of refined models with increased complexity. A fast charging approach with the merits of both multi-objective optimality and online tractability is thereby highly desired.

To fill the aforementioned gaps, this work proposes a machine learning-enabled battery fast charging strategy by combining a validated inner state-aware reduced-order electrochemical model (ROM) with a novel continuous deep reinforcement learning (DRL) algorithm. Three primary contributions are made.

First, the electrochemical modeling is combined with the DRL algorithms, for the first time, to enhance the battery management system. The incorporation of mechanism- and data-driven approaches has never been explored before, and validates to mitigate the unfavorable over-temperature and side reaction based on the inner state perceiving and control.

Second, the cyber-physical and cloud technique has been viewed as the future trend of battery management system. To date, however, there has been no report on the realization of fast charging within such a framework. Motivated by this, the proposed machine learning strategy is performed within a battery-to-cloud framework, for the first time, to mitigate the risk of heavy computing burden while ensure a high adaptability to the charging environment.

Third, the SAC as a state-of-the-art machine learning algorithm is exploited to successfully develop the smart fast charging strategy. The SAC-enabled strategy gives rise to the most stable and reliable charging performance compared to the well-known alternatives in the RL family.

The remainder of this paper is organized as follows. A multi-state-aware ROM is developed and validated in Section 2, which is further used for charging strategy training. The SAC-DRL-based machine learning strategy is elaborated in Section 3. The HiL results and experimental results with practical batteries are discussed in Section 4, while the primary conclusions are drawn in Section 5.

2. Battery modeling

2.1. Full-order electrochemical-thermal model

A classic FOM is established by regarding the LIB as a whole containing three sections, i.e., a porous anode, a cathode, and a separator. In the FOM, a series of partial differential equations are exploited to accurately depict the lithiation, ion diffusions, ion transports, electrochemical reactions, and heat processes during the applications of batteries [22].

The Li-ion diffusions in the electrode and electrolyte can be modeled by the Fick's second law, and then the Li-ion concentrations are depicted as:

$$\frac{\partial c_s}{\partial t} = \frac{D_s}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial c_s}{\partial r} \right) \quad (1)$$

$$\frac{\partial}{\partial t} (\varepsilon_e c_e) = \frac{\partial}{\partial x} \left(D_{eff} \frac{\partial c_e}{\partial x} \right) + \frac{1 - t_c}{F} j_f \quad (2)$$

where D_s , D_{eff} are the effective solid and liquid diffusion coefficient, respectively, t_c is the Li-ion transference number, j_f denotes the volumetric current density.

The local potentials in the electrode and electrolyte can be depicted by the Ohm's law:

$$\frac{\partial}{\partial x} \left(\sigma_{eff} \frac{\partial \varphi_s}{\partial x} \right) - j_f = 0 \quad (3)$$

$$\frac{\partial}{\partial x} \left(\kappa_{eff} \frac{\partial \varphi_e}{\partial x} \right) + \frac{\partial}{\partial x} \left(\kappa_{D,eff} \frac{\partial \ln c_e}{\partial x} \right) + j_f = 0 \quad (4)$$

where σ_{eff} , κ_{eff} are the effective electronic and ionic conductivity in the electrolyte, respectively.

The Li-ion intercalation/deintercalation reactions can be described by the Butler-Volmer equation, which reveals the relationship between the volumetric current and electrode activation polarization over-potential:

$$j_f = a_s i_0 \left[\exp \left(\frac{\alpha_a F}{RT} \eta \right) - \exp \left(- \frac{\alpha_c F}{RT} \eta \right) \right] \quad (5)$$

$$i_0 = k_0 (c_e)^{\alpha_a} (c_s, surf, max)^{\alpha_a} (c_s, surf)^{\alpha_c} \quad (6)$$

where α_a , α_c denote the electrode transfer coefficients in the anode and cathode, respectively.

The thermal dynamics spawned by the electrochemical reactions are described using the thermal conductivity equation:

$$\rho C_p \frac{\partial T}{\partial t} = \frac{\partial}{\partial R} \left(\lambda_R \frac{\partial T}{\partial R} \right) + \frac{\partial}{\partial Z} \left(\lambda_Z \frac{\partial T}{\partial Z} \right) + H \quad (7)$$

$$H = \sigma_{eff} (\nabla \varphi_s)^2 + (\kappa_{eff} \nabla \varphi_e + \kappa_{D,eff} \nabla \ln c_e) \nabla \varphi_e + ITE_n (SoC) \quad (8)$$

where E_n represents the entropy change, C_p , λ_Z , λ_R are the thermal capacity, axial thermal conductivity and radial thermal conductivity, respectively.

2.2. Model reformulation

FOM has an appreciable precision but simultaneously introduces the intractable calculations, which prohibits the design of the controller. Therefore, this section explores a reduced-order electrochemical-thermal model (ROM) to balance the accuracy and computing complexity.

Reformulation 1. : The Li-ion diffusion solution in the solid phase of the Eq. (1) can be approximated by a three-parameter polynomial profile [23]:

$$\frac{dc_{s,avg}^{\pm}(t)}{dt} = -3 \frac{j_n^{\pm}(t)}{R_s^{\pm}} \quad (9)$$

$$\frac{dq_{avg}^{\pm}(t)}{dt} = -30 \frac{D_s^{\pm} q_{avg}^{\pm}(t)}{R_s^{\pm,2}} - \frac{45}{2} \frac{j_n^{\pm}}{R_s^{\pm,2}} \quad (10)$$

$$c_{s,surf}^{\pm}(t) = c_{s,avg}^{\pm}(t) + \frac{R_s^{\pm}}{35D_s^{\pm}} \left(8D_s^{\pm} q_{avg}^{\pm}(t) - j_n^{\pm} \right) \quad (11)$$

where $c_{s,avg}$, $c_{s,surf}$, q_{avg} are the volume-averaged concentration, surface concentration, and volume-averaged concentration flux, respectively, j_n denotes the pore wall flux at the surface of the particle, and $j_n = j_f / (a_s F)$. Superscript \pm denotes the polarity of the electrode.

Reformulation 2. : The Li-ion diffusion solution in the liquid phase of the Eq. (2) can be approximated by the polynomial curve of either the second-order or less [24]. Consequently, the electrolyte concentrations at any position x of both electrodes can be expressed as:

$$c_e^{\pm}(x, t) = \left(\frac{P_1^{\pm} x^2}{2} + P_2^{\pm} \right) f^{\pm}(t) \frac{I(t)}{A^{\pm}} + c_e^{\pm}(x, 0) \quad (12)$$

where P_1 and P_2 are constants depended to the geometry parameters of LIB.

For the charging control problem, the electrolyte concentrations at the boundaries of electrode $x = 0^{\pm}$ are critical. Substituting (12) and $x = 0^{\pm}$ into (2) gives:

$$\frac{dc_e^{\pm}(0^{\pm}, t)}{dt} = \frac{D_e^{\pm} \mu^{\pm}}{\varepsilon_e^{\pm}} [c_e^{\pm}(0^{\pm}, t) - c_{e0}^{\pm}] + \frac{1 - t_c}{\varepsilon_e^{\pm} F} j_f^{\pm}(t) \quad (13)$$

where μ is a constant related to P_1 and P_2 .

Reformulation 3. : The local volumetric transfer current density j_f in each electrode satisfies the following spatial integrals [25]:

$$\int_{0^{\pm}}^{L^{\pm}} j_f^{\pm}(t) dx = \mp \frac{I(t)}{A^{\pm}} = j_{f,avg}^{\pm}(t) L^{\pm} \quad (14)$$

With this assumption, Eqs. (3) and (4) can be derived analytically as a function of the spatial coordinate x :

$$\phi_s^{\pm}(x, t) = \phi_s^{\pm}(0^{\pm}, t) - \frac{I(t)}{A^{\pm} \sigma_{eff}^{\pm}} \left(x - \frac{x^2}{2L^{\pm}} \right) \quad (15)$$

$$\phi_e^{\pm}(x, t) = \phi_e^{\pm}(0^{\pm}, t) + \frac{2(1 - t_c)RT}{F} \ln \left(\frac{c_e^{\pm}(x, t)}{c_e^{\pm}(0^{\pm}, t)} \right) - \frac{x^2}{2\kappa_{eff}^{\pm}} j_f^{\pm}(t) \quad (16)$$

$$-SGN \frac{I(t)}{2} \left(\frac{L^-}{A^- \kappa_{eff}^-} + \frac{2L^{sep}}{A^{sep} \kappa_{eff}^{sep}} + \frac{L^+}{A^+ \kappa_{eff}^+} \right) \quad (16)$$

where $SGN = 0$ when x locates in the anode region, $SGN = 1$ when x locates in the cathode region. Superscript ^{sep} denotes the separator.

Given that the electronic conductivity of the solid phase is much greater than the ionic conductivity of the liquid phase, the potentials at the boundaries of the electrodes $x = 0^{\pm}$ are need and expressed by:

$$\phi_e(t) = \phi_e^+(0^+, t) - \phi_e^-(0^-, t) = \frac{2(1 - t_c)RT}{F} \ln \frac{c_e^+(0^+, t)}{c_e^-(0^-, t)}$$

$$-\frac{I(t)}{2} \left(\frac{L^-}{A^- \kappa_{eff}^-} + \frac{2L^{sep}}{A^{sep} \kappa_{eff}^{sep}} + \frac{L^+}{A^+ \kappa_{eff}^+} \right) \quad (17)$$

Reformulation 4. : Rewriting the Butler-Volmer equation, the electrode activation polarization overpotentials η for the Li-ion intercalation/deintercalation reactions can be derived based on Eq. (5):

$$\eta^{\pm}(t) = \frac{2RT_a(t)}{F} \ln \left(\xi^{\pm}(t) + \sqrt{\xi^{\pm,2}(t) + 1} \right) \quad (18)$$

$$j_0^{\pm}(t) = k^{\pm} \left[\left(c_{s,surf}^{\pm}(t) \right)_{s,surf}^{\pm} (0^{\pm}, t) \right]^{0.5} \quad (19)$$

where $\xi^{\pm}(t) = j_f^{\pm}(t) / [2a_s^{\pm} j_0^{\pm}(t)]$.

Reformulation 5. : Overlooking the axial distribution of LIB, a series of thermal resistances and heat capacities are employed to approximate the processes of the conduction and convection, and thus the internal and surface temperature can be characterized by the radial energy balance equations [26]:

$$\frac{dT_s(t)}{dt} = \frac{T_c(t) - T_s(t)}{R_c C_s} + \frac{T_f - T_s(t)}{R_u C_s} \quad (20)$$

$$\frac{dT_c(t)}{dt} = \frac{T_s(t) - T_c(t)}{R_c C_c} + \frac{H(t)}{C_c} \quad (21)$$

$$H(t) = I(t)[E(t) - V_i(t)] + I(t)[T_a(t) + 273]E_n(\text{SoC}, t) \quad (22)$$

$$T_a(t) = \frac{T_c(t) + T_s(t)}{2} \quad (23)$$

where T_s , T_c and T_f are the battery surface, internal and ambient temperature, respectively, R_c and R_u are thermal resistances related to the heat conduction inside the battery and the convection at the battery surface, C_c and C_s are equivalent thermal capacitances of the battery core and surface, H is the heat generation rate, including the ohmic heat, polarization heat and entropic heat.

In line with the aforementioned simplifications, the electromotive force (EMF), SoC and the terminal voltage of LIB can be calculated as:

$$E(t) = U^+ \left(\frac{c_{s,surf}^+(t)}{c_{s,max}^+} \right) - U^- \left(\frac{c_{s,surf}^-(t)}{c_{s,max}^-} \right) \quad (24)$$

$$\text{SoC}(t) = \frac{c_{s,avg}^-(t) - c_{s,min}^-}{c_{s,max}^-} \quad (25)$$

$$V_i(t) = E(t) + \phi_e(t) + \eta^+(t) - \eta^-(t) - R_f I(t) \quad (26)$$

where $U^{\pm}(\cdot)$ is a function of solid particle surface concentration, which indicates the open circuit potential of the corresponding electrode, R_f is the SEI film resistance.

Finally, to capture the impact of the high-rate charging on the battery aging, the anodic overpotential of side reaction is taken into account in the model [27]. The side-reaction overpotential η_{side} is given by:

$$\eta_{side}(t) = \eta^-(t) + U^-(t) - U_{side} \quad (27)$$

where U_{side} denotes the equilibrium potential of the side reaction.

2.3. Model validation

The A123 commercial LFP cell is charged and discharged under different conditions using the Arbin testing system to benchmark the ROM elucidated herein. The parameters involved in the simulations are indicated in Table 1. Specifically, the battery is located in a thermal chamber to control the ambient temperature at 25 °C. Three thermocouples are attached along the axial direction to record the surface temperature of LIB. The measured terminal voltage and surface temperature are compared with their predicted results to validate the model, as shown in Fig. 1. Explicitly, it turns out that the ROM can simulate the electrical and thermal dynamics of LIB at the different C-rates. The mean absolute errors (MAE) of the voltage and temperature under constant current conditions are 0.023 V and 0.29 °C, respectively,

Table 1
Parameters of the ROM.

Parameters	Symbol [unit]	Negative electrode	Separator	Positive electrode
Electrode thickness	L [μm]	70 ^a	34 ^a	25 ^a
Electrode plate area	A [m^2]	0.18 ^a	0.18 ^a	0.18 ^a
Particle radius	R_p [μm]	5 ^a	–	0.0365 ^a
Volume fraction of solid phase	ϵ_s [-]	0.56 ^b	0.55 ^a	0.432 ^b
Volume fraction of electrolyte	ϵ_e [-]	0.36 ^b	0.45 ^a	0.426 ^b
Maximum Li-ion concentration in solid	$c_{s,\text{max}}$ [$\text{mol}\cdot\text{m}^{-3}$]	30,555 ^a	–	51,410 ^a
Stoichiometry at 0% SoC	x_0, y_0 [-]	0.0306 ^b	–	0.9586 ^b
Stoichiometry at 100% SoC	x_{100}, y_{100} [-]	0.6873 ^c	–	0.4719 ^c
Maximum stoichiometry range	$\Delta x, \Delta y$ [-]	0.6567 ^b	–	0.4867 ^b
Initial Li-ion concentration in electrolyte	c_{e0} [$\text{mol}\cdot\text{m}^{-3}$]	1000 ^a	–	–
Li-ion diffusion coefficient in solid	D_s [$\text{m}^2\cdot\text{s}^{-1}$]	9.891×10^{-14} ^b	–	1.28×10^{-18} ^b
Li-ion effective diffusion coefficient in electrolyte	D_{eff} [$\text{m}^2\cdot\text{s}^{-1}$]	1.2×10^{-10} ^b	–	–
Ionic conductivity	κ_e [$\text{S}\cdot\text{m}^{-1}$]	0.93 ^b	–	–
Bruggeman coefficient	$Brugg$	1.5 ^a	–	–
Reaction rate coefficient	k_0 [$\text{A}\cdot\text{m}^{-2}$]	3.38×10^{-11} ^b	–	8.8×10^{-12} ^b
Li-ion transference number	t_c [-]	0.363 ^a	–	–
Charge transfer coefficient	α_w, α_c [-]	0.5 ^a	–	0.5 ^a
Equilibrium potential of lithium plating	U_{side} [V]	0 ^a	–	–
Film resistance of SEI layer	R_f [Ω]	0.0012 ^b	–	–
Faraday constant	iF [$\text{C}\cdot\text{mol}^{-1}$]	96,485 ^a	–	–
Universal gas constant	R [$\text{J}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$]	8.314 ^a	–	–
Internal heat conduction resistance	ii) R_k [$\text{K}\cdot\text{W}^{-1}$]	9.52 ^b	–	–
Surface heat convection resistance	R_{li} [$\text{K}\cdot\text{W}^{-1}$]	12.55 ^b	–	–
Core heat capacity	C_c [$\text{J}\cdot\text{K}^{-1}$]	87.69 ^b	–	–
Surface heat capacity	C_s [$\text{J}\cdot\text{K}^{-1}$]	4.28 ^b	–	–
Negative electrode equilibrium potential (Graphite) ^a :				
$U_n = 0.6379 + 0.5416\exp(-305.5309x) + 0.044\tanh\left(-\frac{(x-0.1958)}{0.1088}\right) - 0.1978\tanh\left(\frac{(x-1.0571)}{0.0854}\right) - 0.6875\tanh\left(\frac{(x+0.0117)}{0.0529}\right) - 0.0175\tanh\left(\frac{(x-0.5692)}{0.0875}\right)$				
Positive electrode equilibrium potential (LFP) ^a :				
$U_p = 3.4323 - 0.8428\exp(-80.2493(1-y)^{1.3198}) - 3.2474 \times 10^{-6}\exp(20.2645(1-y)^{3.8003}) + 3.2482 \times 10^{-6}\exp(20.2646(1-y)^{3.7995})$				

^a From reference [28,29].

^b Identified using PSO.

^c Calculated.

and the corresponding errors under dynamic stress test (DST) condition are 0.027 V and 0.33 °C.

The internal parameters and states are essential for the regulation of charging process. Hence, the developed model is expected to be accurate for inner process description. Specifically, the internal temperature is evaluated using the cell embedded the thermal sensor to validate the performance of thermal sub-model. In detail, a distributed optical fiber sensor is implanted into the core of cylindrical cell, and the Rayleigh scattering principle is then used to achieve high precision sensing of the internal temperature of the battery. The relevant embedded thermal monitoring technique has been elaborated in our previous works [30, 31], and thus are not further explained herein. With the implanted sensors, the internal temperatures under constant current conditions with different rates are compared with results of the developed ROM. The MAE and rooted mean squared errors (RMSE) are shown in Table 2. It is explicit that the modelled internal temperatures are in good agreement with the measured benchmarks, validating the high fidelity of the model regarding perceiving the internal temperature of LIB.

It is noted that as for the side reaction sub-model, there is no sensor that can directly measure and validate the inner side reaction overpotential of battery at present. In this paper, the test data obtained from the half-cell with the lithium reference electrode is used to correct the open circuit potential of electrode in the electrochemical model. Subsequently, the overpotential is calculated using widely well-recognized governing equations. Therefore, the side reaction overpotential can be considered to be approaching the ground truth although not validated directly.

3. SAC-based fast charging strategy

On the premise of the ROM developed in Section II, this section goes further to propose a LIB fast charging strategy with safety and degradation consciousness based on the SAC-DRL algorithm.

The charging aims to reach the pre-determined SoC as fast as possible, while the risk of over-temperature, over-/under-voltage, and side reactions can be well mitigated. To be brief, a multi-physics-constrained charging problem is formulated as:

$$\begin{aligned} \min_{I(t)} \quad & \int_{t_0}^{t_f} |SoC_{\text{tar}} - SoC(t)| dt \\ \text{s.t.} \quad & 0 \leq I(t) \leq I_{\text{max}} \\ & V_{t,\text{min}} \leq V(t) \leq V_{t,\text{max}} \\ & T_{\text{min}} \leq T_c(t) \leq T_{\text{max}} \\ & \eta_{\text{side}}(t) \geq 0 \end{aligned} \quad (28)$$

In particular, since the side reactions like the lithium plating and SEI growth are more likely to arise when $\eta_{\text{side}} < 0$, it is essential to adjust the charging current adaptively to avoid the situation where the side reaction overpotential drops to zero [32].

3.1. State observer

The formulated charging optimization method relies on the active constraints on multiple immeasurable states. Therefore, a state observer is designed herein to estimate these variables, including T_c , SoC and η_{side} . The ROM equations aforementioned are rewritten as a state-space form, in which the input is $u = I$, the state vector is $x = [c_{s,\text{avg}}^-, c_{s,\text{avg}}^+, q_{\text{avg}}^-, q_{\text{avg}}^+, c_e^+(0^+) - c_{e0}^+, c_e^-(0^-) - c_{e0}^-, T_s, T_c]^T$, the measurable outputs vector is $z = [V_t, T_s]^T$, and the required outputs vector is $y = [SoC, V_t, T_c, \eta_{\text{side}}]^T$. Given the nonlinearity of the ROM, an extended Kalman filter (EKF) approach is utilized for observation, with the detailed procedures generalized in Table 3. where \mathbf{W}_k and \mathbf{V}_k are the covariance matrix of the process noises and measurement noises evaluated at the k -th time step.

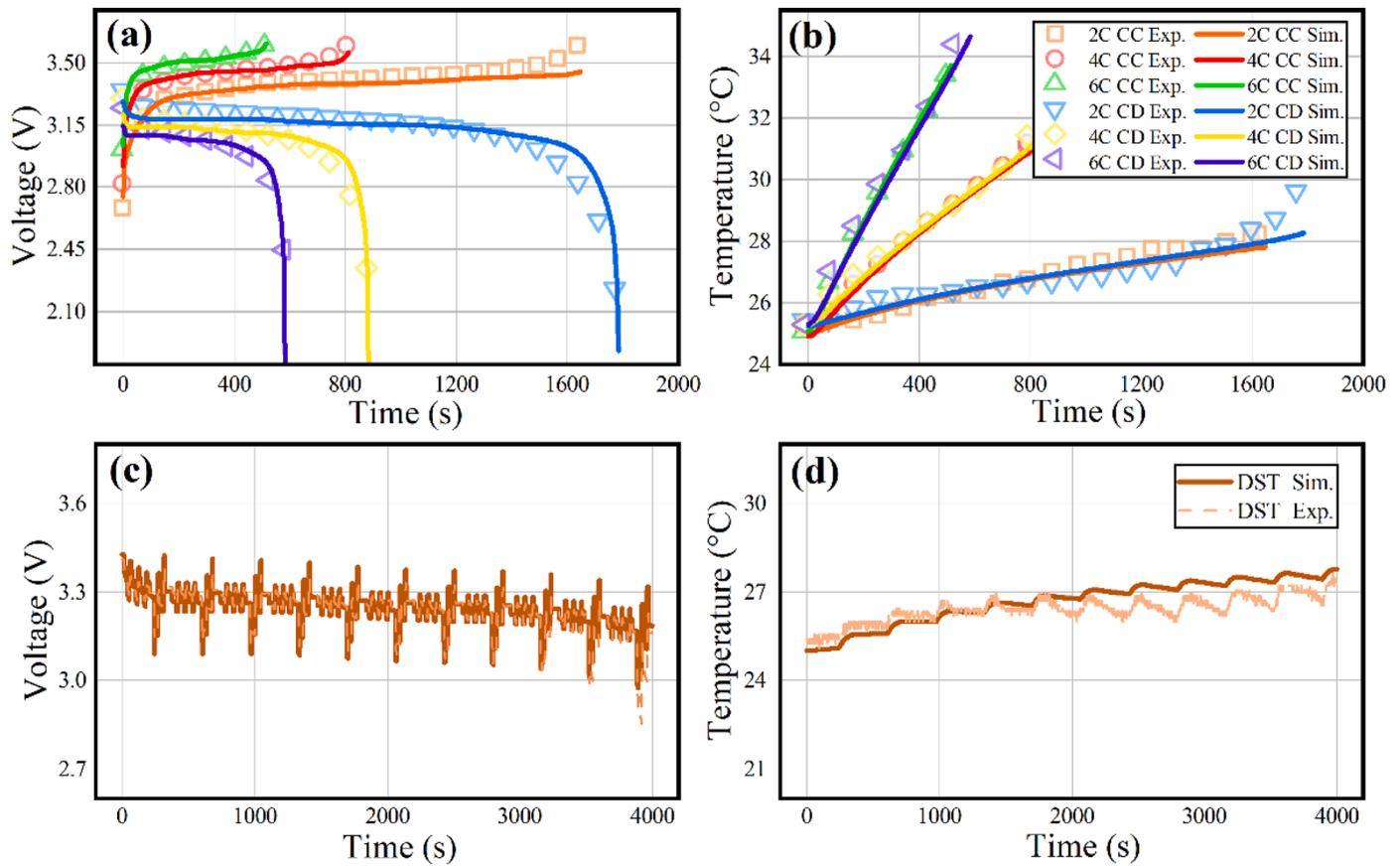


Fig. 1. Validation results of battery modeling: (a) terminal voltage and (b) surface temperature of constant current conditions, (c) terminal voltage and (d) surface temperature of DST condition.

Table 2
MAE and RMSE of internal temperatures under different rates.

	Internal temperature (°C)	
	Low rate	High rate
MAE	0.12	0.34
RMSE	0.18	0.42

3.2. SAC algorithm

Different from the general actor-critic algorithms, the SAC incorporates the maximum entropy in the reward, aiming to increase the exploration abilities and robustness of the agent. Accordingly, the soft Q-function is re-determined based on the Bellman iteration [33]:

$$Q(s_t, a_t) = r(s_t, a_t) + \gamma E_{s_{t+1}, a_{t+1}} [Q(s_{t+1}, a_{t+1}) - \text{alog}(\pi(a_{t+1}|s_{t+1}))] \quad (29)$$

where $r(s_t, a_t)$ represents the single-step reward given by the environment with respect to step t , γ denotes the discount factor, α is the weighting factor of the entropy term, $\pi(\cdot)$ is the stochastic policy, $E[\cdot]$ is the mathematical expectation.

For the soft policy improvement, the policy is updated in the light of the principle of minimizing the KL-divergence:

$$\pi_{new} = \underset{\pi' \in \Pi}{\text{argmin}} D_{KL} \left(\pi'(\cdot|s_t) \parallel \frac{\exp(\frac{1}{\alpha} Q^{\text{old}}(s_t, \cdot))}{Z^{\text{old}}(s_t)} \right) \quad (30)$$

where Π is a set of the potential policies, $Z^{\text{old}}(s_t)$ is the partition function

Table 3
Procedures of the EKF.

1. Initialization: $\hat{x}_{k-1}, P_{k-1}, W_k, V_k$
2. Definition: $A_k = \frac{\partial f(x_{k-1}, u_{k-1})}{\partial x}$ and $C_k = \frac{\partial h(x_k, u_k)}{\partial x}$
3. **repeat**
4. State prediction: $\hat{x}_k^- = f(x_{k-1}, u_{k-1})$
5. Covariance prediction: $P_k^- = A_k P_{k-1} A_k^T + W_k$
6. Kalman gain: $K_k = P_k^- C_k^T (C_k P_k^- C_k^T + V_k)^{-1}$
7. State correction: $\hat{x}_k = \hat{x}_k^- + K_k (z_k - h(\hat{x}_k^-, u_k))$
8. Covariance correction: $P_k = (I - K_k C_k) P_k^-$
9. **until stop**

to normalize the distribution, and makes no any difference on the policy gradient.

In general, deep neural networks (DNNs) are used to approximate the soft Q-function and the policy π , where the network parameters are θ and φ , respectively. The soft Q-function network Q_θ directly outputs a Q value, yet the output of the policy network π_φ is generally defined by a Gaussian distribution, including the mean and standard deviation. Meanwhile, the target network Q_θ^- is congregated to boost the stability of training, which possesses the same network structures and initial weights with Q_θ . Note that the parameters of each DNN have to be updated by the successive training to guarantee the accuracy of the approximation.

The parameters of the soft Q-function networks can be learned by minimizing the soft Bellman residual:

$$J_Q(\theta) = E_{(s_t, a_t, s_{t+1}) \sim D, \tilde{a}_{t+1} \sim \pi_\theta} \left[\frac{1}{2} \left(Q_\theta(s_t, a_t) - \left(r(s_t, a_t) + \gamma \left(Q_{\tilde{a}}(s_{t+1}, \tilde{a}_{t+1}) - \alpha \log(\pi_\theta(\tilde{a}_{t+1} | s_{t+1})) \right) \right) \right)^2 \right] \quad (31)$$

where D denotes the experience pool, which accommodates a mass of transitions (s_t, a_t, s_{t+1}, r_t) . It is worth noting that a_t denotes the action actually performed from the experience pool D , while \tilde{a}_t is the sample from $\pi_\theta(\cdot | s_t)$ representing a potential action.

The parameters of the policy network can be trained to minimize the expected KL-divergence in Eq. (30):

$$J_\pi(\varphi) = E_{s_t \sim D, \tilde{a}_t \sim \pi_\varphi} \left[\log \pi_\varphi(\tilde{a}_t | s_t) - \frac{1}{\alpha} Q_\theta(s_t, \tilde{a}_t) + \log Z(s_t) \right] \quad (32)$$

To achieve the minimum of J_π by the gradient descent, the reparameterization trick is applied to ensure the backward propagation of the gradients. Hence, the \tilde{a}_t in Eq. (32) is replaced with:

$$\tilde{a}_{\varphi,t} = f_\varphi^\mu(s_t) + \varepsilon_t \odot f_\varphi^\sigma(s_t) \quad (33)$$

where $f_\varphi^\mu(s_t)$ and $f_\varphi^\sigma(s_t)$ correspond to the mean and standard deviation of the policy network π_φ , ε_t represents the noise sampled from the standard normal distribution.

A soft updating strategy is adopted for the parameters of the target network:

$$\bar{\theta} \leftarrow \tau \bar{\theta} + (1 - \tau) \theta \quad (34)$$

where τ is the factor of the soft update.

The weighting factor α is a non-negative hyperparameter, which value needs to be tuned in time to adapt to the training process. In general, it can be updated by minimizing the following objective function:

$$J(\alpha) = E_{\tilde{a}_t \sim \pi_\varphi} \left[-\alpha \log \pi_\varphi(\tilde{a}_t | s_t) - \alpha \mathcal{H}_0 \right] \quad (35)$$

where \mathcal{H}_0 denotes the target entropy, whose value is equal to $-\dim(a_t)$.

3.3. SAC-based charging control

To solve the formulated optimization problem (28) with the SAC algorithm, the state and action variables are defined as:

$$\begin{cases} s_t = \{f_{nor}[SoC(t)], f_{nor}[V_t(t)], f_{nor}[T_c(t)], f_{nor}[\eta_{side}(t)]\} \\ a_t = \{f_{nor}[I(t)] \mid I(t) \in [0, I_{max}]\} \end{cases} \quad (36)$$

where $f_{nor}[\cdot]$ is used to transform the variables into a unified range of $[-1, 1]$, $I(t)$ denotes the charging current. Particularly, $V_t(t)$ can be measured directly, while $SoC(t)$, $T_c(t)$ and $\eta_{side}(t)$ should be online estimated using the state observer developed in Section III-B.

Through the continuous interactions with the battery environment, the agent of SAC can optimize the charging by maximizing the long-term accumulative reward. A non-sparse and real-time reward function is formulated as:

$$r(s_t, a_t) = -\lambda [\omega_1 C_{soc}(t) + \omega_2 C_{volt}(t) + \omega_3 C_{temp}(t) + \omega_4 C_{side}(t) + \omega_5 C_{smooth}(t)] \quad (37)$$

where λ is the reward scaling factor in ensuring a relatively minor difference in both the reverse gradient of the maximum entropy term and the cumulative reward term. ω_i ($i = 1 \sim 5$) are the weights to evaluate the level of concern for each constraint.

The first term denotes the objective function to be minimized:

$$C_{soc}(t) = |SoC_{tar} - SoC(t)| \quad (38)$$

$C_{volt}(t)$, $C_{heat}(t)$ and $C_{side}(t)$ are arranged as a similar form to address

the constraints in Eq. (28) on voltage, temperature and side reaction, respectively. For simplicity, $C_{volt}(t)$, $C_{heat}(t)$ and $C_{side}(t)$ can be expressed as:

$$\begin{aligned} C_{volt}(t) &= \begin{cases} 0, & V_{t,min} \leq V_t(t) \leq V_{t,max} \\ 1, & \text{other} \end{cases} \\ C_{heat}(t) &= \begin{cases} 0, & T_c(t) \leq T_{max} \\ 1, & T_c(t) > T_{max} \end{cases} \\ C_{side}(t) &= \begin{cases} 0, & \eta_{side}(t) \geq 0 \\ 1, & \eta_{side}(t) < 0 \end{cases} \end{aligned} \quad (39)$$

Especially, an additional term $C_{smooth}(t)$ is introduced into the reward to suppress the heavy fluctuation of current during the charging:

$$C_{smooth}(t) = |I(t) - I(t-1)| \quad (40)$$

With the aforementioned definition of the reward function, state and action, an SAC-based optimal charging policy can be trained for further online utilization. The soft Q -function network Q_θ is comprised of two fully connected layers with 256 nodes, followed by an output layer containing the $\tanh(\cdot)$ activation function. The double Q -function networks are employed simultaneously to enhance the robustness. The policy network π_φ is provided with the same hidden layer structure as the soft Q -function network. Nevertheless, the output layer obtains the mean and standard deviation of a Gaussian distribution simultaneously. Beyond that, an Adam optimizer is used and the detailed hyperparameters can be found in Table 4. To exemplify this, the detailed procedure of training is illustrated in Fig. 2 and Table 5, while the practical use of the trained charging policy is depicted in Fig. 3 and Table 6.

3.4. Implementation within battery-to-cloud framework

Attributed to the separable feature of offline training and real-time charging control, the SAC-based fast charging strategy is merged into a cloud-based management system [34]. As shown in Fig. 4, the training of strategies is performed in a cloud server while the real-time decision is applied to the battery. In the cloud, the calibrated battery model is used as the environment, which provides massive data covering a broad range of both favorable and abusive charging scenarios. Relying on the generated dataset, the optimal charging policy can be trained by the efficient interface between the agent and the environment. It is worth noting that the battery model on the cloud will be updated periodically using the new data to ensure that the charging strategies can be adapted to the aging of the battery. Once the training is accomplished, the determined charging policy is downloaded to the onboard controller for optimal charging. Empowered by this, the computational loads on the battery controller can be decreased significantly.

Table 4
Hyperparameters of SAC.

Hyperparameters	Value	Description
N_D	400,000	Size of experience pool
N_B	128	Size of minibatch
M	600	Maximum training episodes
lr	0.0002	Learning rate
τ	0.005	Soft update factor
γ	0.99	Discount factor

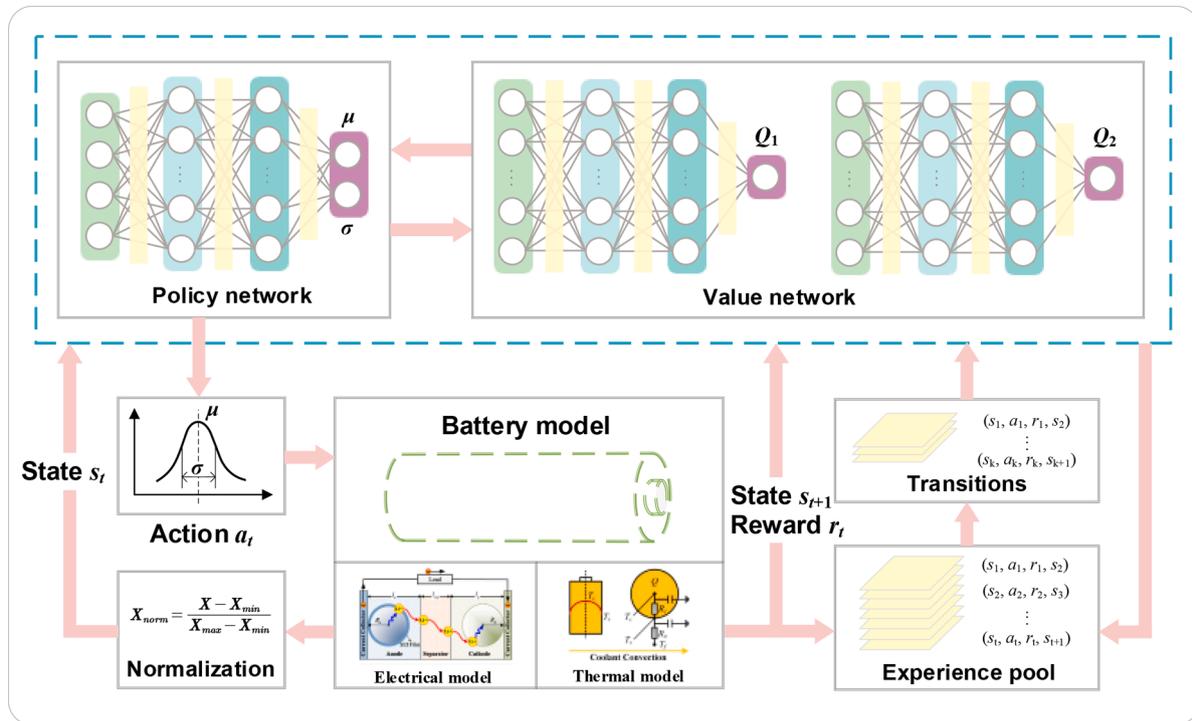


Fig. 2. Training procedure of fast charging strategy based on SAC.

Table 5

Training procedure of fast charging strategy.

1. Initialize policy parameters φ , double soft Q-function parameters $\theta_1, \theta_2, \bar{\theta}_1, \bar{\theta}_2$, and an empty experience pool D
2. **repeat**
3. For time t , observe state s_t
4. Select action $a_t \sim \pi_\varphi(\cdot|s_t)$, and convert action a_t into the real charging current I_t
5. Execute I_t in the battery model environment
6. Observe the next state s_{t+1} , reward r_t , and d_t to indicate s_{t+1} whether the charging is completed
7. Save a transition $(s_t, a_t, r_t, s_{t+1}, d_t)$ in the experience pool D
8. If the charging is completed, reset the environment and the initial state with a certain randomness
9. if the experience pool is full, **then**
10. **for** j in range (M) **do**
11. Randomly sample a batch of transitions, $B = [(s_b, a_b, r_b, s_{b+1}, d_b)]$, from D
12. Update Q-functions with Eq. (31)
13. **if** $j \bmod M = 0$ **then**
14. Update the policy with Eq. (32)
15. Update the weighting factor α with Eq. (35)
16. Update the target networks with Eq. (34)
17. **end if**
18. **end for**
19. **end if**
20. **until** convergence

4. Results and discussion

4.1. Results of policy training

Following the fundamental framework in Fig. 4, a primary collaboration platform is established for the theoretical validation by combining a workstation and hardware-in-the-loop (HiL) system. In detail, a workstation (Intel Xeon Sliver, 32GB RAM, Tesla T4) is used as the cloud server to update and store the battery model and to train the charging policy. In the meantime, a low-cost embedded controller (RapidECU-U4B, 32 bit, 64Kb EEPROM) is used as an onboard controller to control the charging process in real-time. The CAN bus is used for communication between the cloud and the battery.

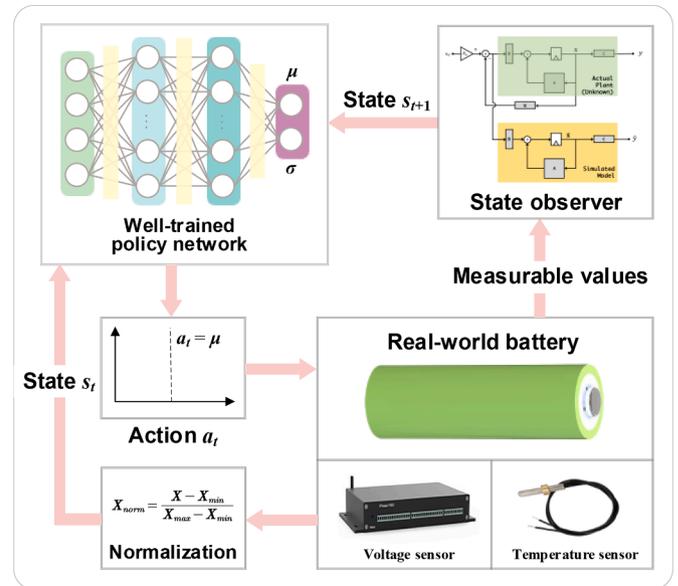


Fig. 3. Fast charging strategy based on SAC.

Table 6

Procedures of the fast charging strategy.

1. Download the policy network parameters φ to the charging control unit
2. Set the charging target
3. **repeat**
4. For time t , obtain the present state s_t
5. Select action $a_t = \text{mean}[\pi_\varphi(\cdot|s_t)]$, and convert the action a_t into the real charging current I_t
6. Execute I_t in the real-world environment
7. Obtain the next state s_{t+1} and d_t to indicate s_{t+1} , and observe whether it satisfies the charging target
8. **until** $d_t = 1$, then the charging is completed

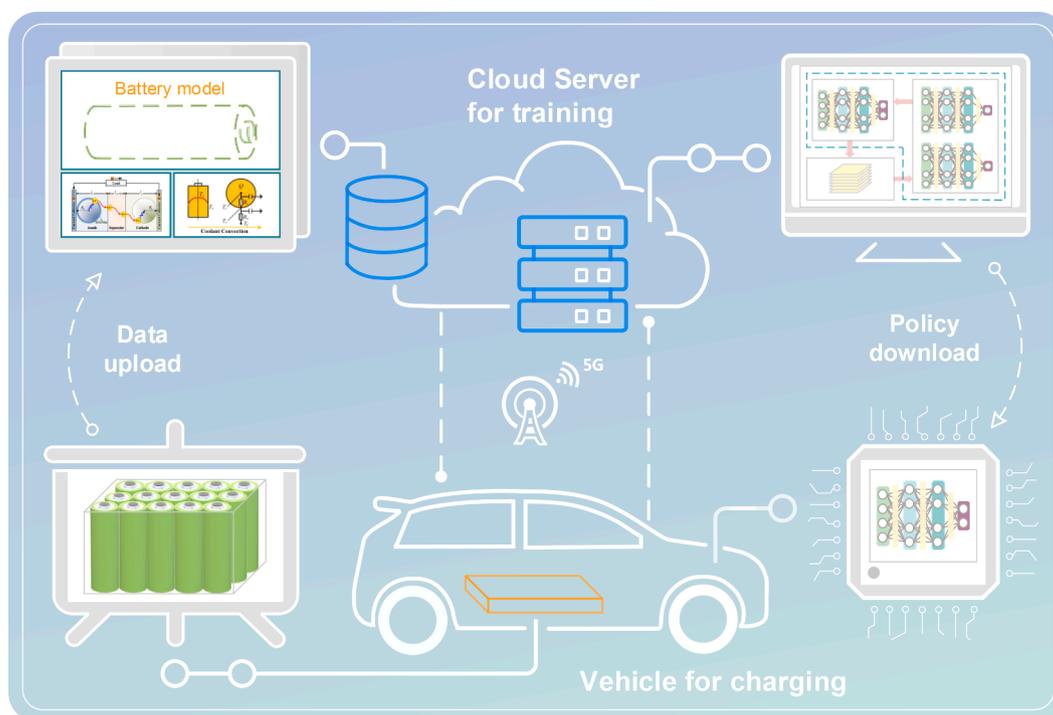


Fig. 4. The framework of the SAC-based charging strategy enabled by cloud computing.

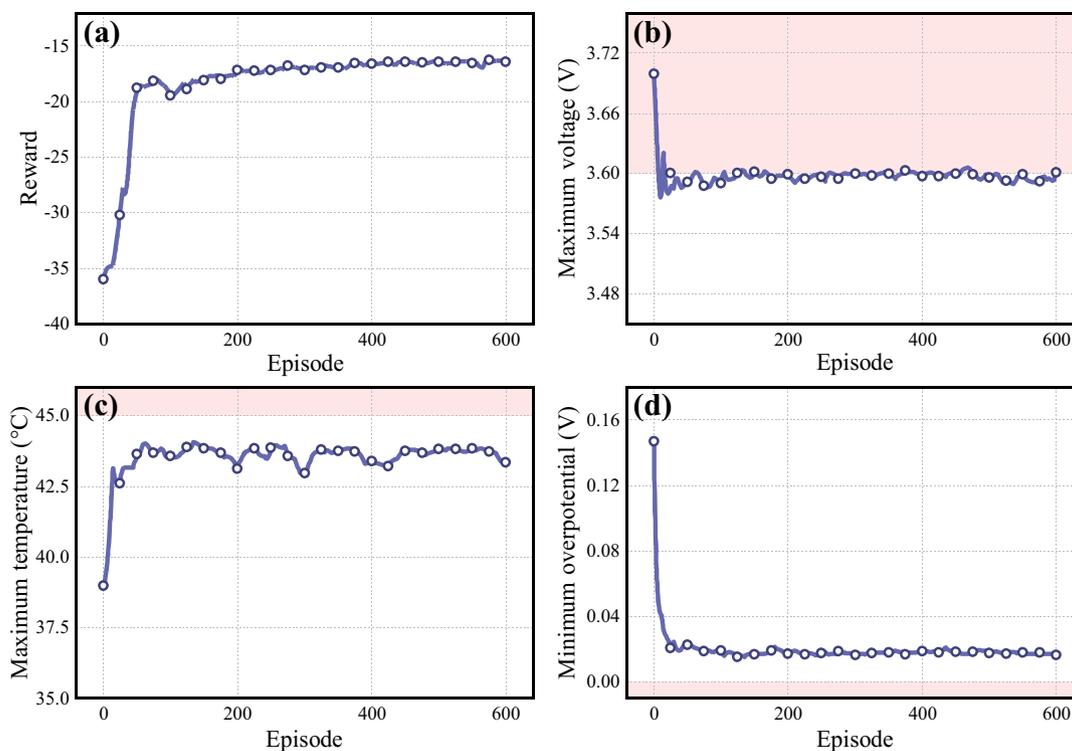


Fig. 5. Evaluation indicators of each episode during the training: (a) reward, (b) maximum voltage during, (c) maximum temperature, (d) minimum side reaction overpotential during charging.

The training of the charging strategy is first performed on the cloud server platform, where the network parameters are successively updated along with the learning proceeds. To validate the effectiveness of learning, the saved each generation of the networks are performed to interact with the same environment, and the key evaluation indicators are plotted in Fig. 5(a) to assess the convergence performance.

Apparently, the reward increases rapidly along with the growing episodes at the beginning stage, indicating efficient learning of the optimal charging strategy through the interaction with the environment. Starting from the ~200th episode, the reward shows a slightly upward and gradually stabilized trend, which portends the convergence of training. To verify this conjecture, the evolutions of the maximum voltage,

maximum temperature, and minimum side reaction overpotential given by the strategies are illustrated in Fig. 5 (b-d). It is not surprising that the three evaluation indicators are fairly close to their pre-set limits at the end of training. The observed results suggest that the learned policy has promised optimality regarding the overall objective without violating the defined electrical and thermal limitations.

4.2. Hardware-in-the-loop validation

The trained policy is further implemented in a real-time environment to validate the performance. Specifically, the trained policy on the cloud server is downloaded to an onboard controller to carry out the HiL experiment. The HiL system is shown in Fig. 6, where the developed ROM is implanted in a real-time simulator to form a “virtual battery” to replace the real battery in Fig. 3. This can rule out the uncertainties raised from the model mismatch and thus contributes to evaluating the trained strategy from the pure theoretical perspective.

The charging control results under different initial conditions are shown in Fig. 7. It is observed that the permissible maximum charging current is selected automatically to accelerate the charging at the beginning of charging since the voltage, core temperature, and side reaction overpotential are all far away from the safety threshold. As the charging proceeds, the constraining effects of different variables can be observed. First, the core temperature or side reaction overpotential approach the pre-specified threshold, which potentially induces the safety risk and quick degradation of the battery. In response to this, the charging current is decreased so that the temperature or overpotential can always be clamped at the safety boundary without violation. This charging mode is maintained until the constant voltage charging stage, where the current is reduced further to constrain the terminal voltage within the safety limit. This is equivalent to the CV stage of the traditional CCCV charging method. It is hence validated that the fastest charging trajectory mitigating the abusive operations has been found. This is rooted in the awarding of quick SoC increase and the effective constraint on multiple safety- and degradation-relevant states.

It is also noted that the proposed method is robust to the change of charging environment, witnessed by the consistent charging performance at different initial SoCs and ambient temperatures. This can be

explained by two attributes of the employed SAC algorithms. First, certain randomness has been incorporated into the environmental initialization process during the policy training. This enriches the diversity of the traversed states and contributes to broadening the generality of the trained agent. Second, the policy network has the merit of self-adaptability so that it can tolerate the unknown inputs to some extent. The aforementioned encouraging results suggest that the proposed strategy enjoys high optimality and good generality, which is promising for the practical use of fast charging.

4.3. Experimental validation with commercial batteries

Considering that the modeling error during agent learning can degrade the control performance in practical applications, the proposed charging strategy is further evaluated experimentally with commercial batteries in this section. Moreover, the proposed strategy was compared with some state-of-the-art benchmarks, such as the rule-based and model-based strategies.

The CCCV charging protocol as the most commonly used rule-based method is first selected as a baseline to compare with the proposed strategy. The charging results using the 4C CCCV, 6C CCCV charging protocols and the proposed strategy are illustrated comparatively in Fig. 8. To evaluate the charging rapidity, the charging time to different SoCs, i.e., 80%, 90% and 100%, are summarized in Table 6. The conflict between the charging speed and the physical limit compliance can be observed easily from the comparative plots. It is shown that the core temperature and side reaction overpotential are always within a reasonable safety range by using the 4C CCCV charging approach. This is consistent with the battery dataset from the manufacturer, where the charging current is recommended to be no larger than 4C to ensure battery safety and longevity. Nevertheless, the charging time of 4C CCCV strategy is significantly longer than in the other two cases.

The consumed charging time is shortened as the charging current climbs up to 6C. However, the accelerated charging speed is at the expense of over-temperature and the occurrence of subzero overpotential, which is unfavorable to the safety and longevity of LIB. It is shown that the core temperature overshoots the defined threshold to around 46 °C during the 6C charging and the subsequent early stage of

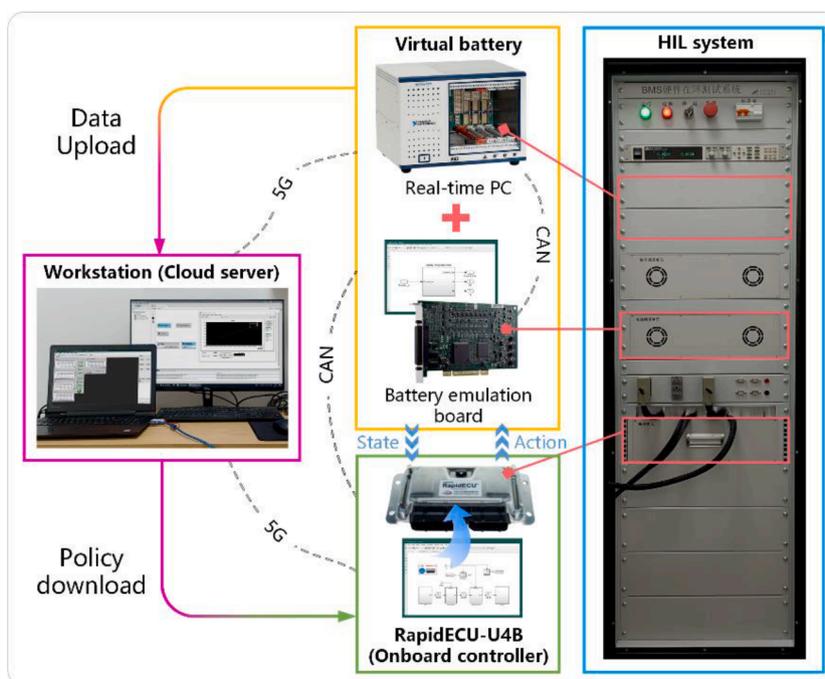


Fig. 6. Diagram of the HiL testing system.

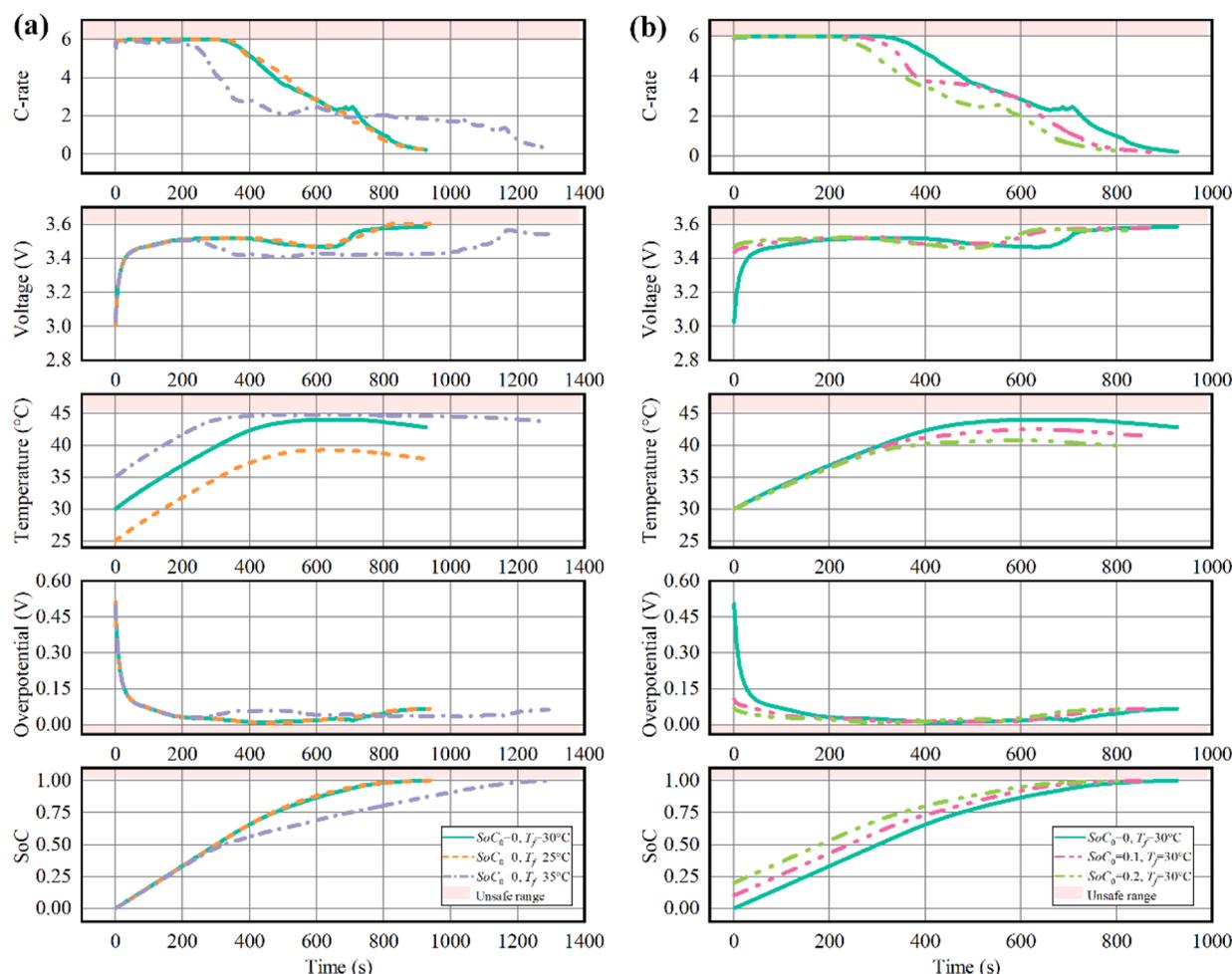


Fig. 7. Validation results of the proposed strategy with HiL experiments.

CV charging. This situation can be much worse if the batteries are used in the tropical zone, where the surrounding temperature in the battery pack can be easily over 40 °C. Moreover, the side reaction overpotential shows to drop below zero from ~400 s to ~600 s during the charging. The subzero overpotential is known as an important contributor to the lithium plating, which is highly dangerous considering its potential to trigger the quick and irreversible degradation, internal short circuit and thermal runaway [35]. It is hence validated that the CCCV charging strategy is non-optimal from the viewpoint of multi-objective fulfillment.

By comparison, the proposed strategy can optimally adjust the charging current depending on the practical triggering condition of the physical limits. Consistent with the HiL results, the whole charging process in present case is divided into three distinguished stages by two distinct turning points, i.e., (1) constant current charging, (2) constant overpotential charging, and (3) constant voltage charging. Each turning point indicates the change of the governing constraining factor. Illustratively, the current first acts as the governing constraint, due to the considerable safety margin of the temperature, voltage and overpotential at the start of charging, and thus the charging curve exhibits a constant-maximum-current mode. Subsequently, the side reaction overpotential begins to take over the governing constraint after the first turning point at around 300 s, and the charging current is down-regulated so that the overpotential can be clamp beyond the 0 V to prohibit the lithium plating. Meanwhile, the current decreases in the slowest way possible, that is, maintaining the potential at zero. Similarly, the governing constraint changes to the voltage at around 650 s at the final of charging, and the current is further reduced, which is

equivalent to the traditional CV charging. It is noted that in this case, the drop in current due to constant overpotential also ensures that the temperature is indirectly controlled below 45 °C, otherwise there will be an additional constant temperature charging phase caused by the temperature taking over the governing constraint.

It is observed that the constrained temperature and overpotential are controlled around the pre-defined thresholds closely, while slight deviations still exist. This can be explained by the modeling mismatch and state estimation error in practical applications, which can inevitably deteriorate the proposed strategy. By adjusting the charging current smartly, the proposed strategy ensures the fulfillment of all the considered physical limits, while meanwhile, the charging speed is appealing considering the comparative time consumption with the 6C CCCV charging. In summary, the proposed charging strategy can be regarded as an optimized balance between the 4C CCCV and 6C CCCV charging strategies.

In addition, the data-driven-based MCC–CV [9] and MPC strategy [18] as the state-of-the-art rule-based and model-based strategies, respectively, have been selected for comparison with the proposed strategy. Specifically, the MCC–CV strategy with the longest cycle life is selected as a delegate one among the candidates for the same charging speed. The comparative results are shown in Fig. 9 and Table 7. With respect to the charging safety, the voltage, temperature and side reaction overpotential are controlled well within the imposed thresholds for all three strategies. It is hence foreseeable that three strategies all meet the requirements of long-term safe use. However, the charging speed are distinguished for different strategies. It is shown that the charging time given by the MPC and proposed strategy are similar and much faster

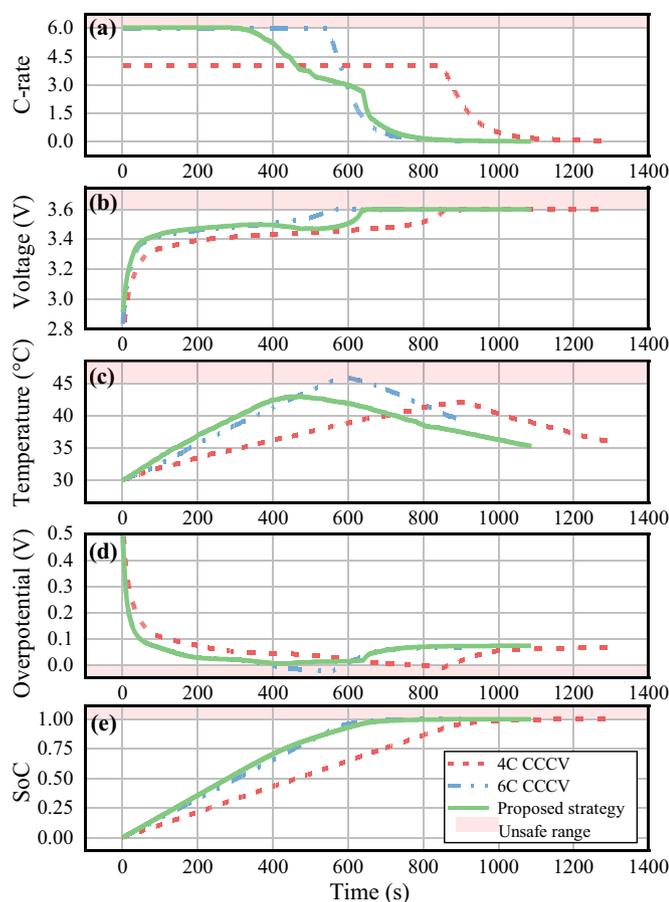


Fig. 8. Comparison of the proposed strategy with the common CCCV strategies: (a) current, (b) terminal voltage, (c) core temperature, (d) side reaction overpotential, (e) SoC.

than the MCC–CV. Notably, the rule-based MCC–CV strategy gives a non-optimal conservative solution, while the former two strategies exhibit the fastest charging while keep the states of interest well within the physical limitations.

Despite the similar optimality, the computing complexity of proposed strategy is much lower than the MPC. To justify this, the computing times occupied per step are used to measure the computational complexity of different strategies. Results show that the computing times are 526 μ s and 216 ms, respectively, for performing the two strategies. Explicitly, the computing resource occupation of the proposed strategy is two orders of magnitude lower than that of the MPC. This is rooted in the fact that the trained strategy involves only easy matrix manipulation for the neural networks, while the MPC performs the computationally demanding moving-horizon optimization at every step. Therefore, provided the comparable optimality, the proposed strategy is more tractable than the MPC in real-world applications.

Long-term cycling experiments are performed to validate the health-conscious performance of the proposed strategy. In particular, the 6C CCCV protocol and MPC strategy are used to compare with the proposed strategy given their similar charging times, which ensure a fair comparison of the anti-aging performance. The capacity fades using different strategies are shown in Fig. 10. From the perspective of aging, the proposed strategy and MPC show similar and explicitly lower capacity decline rates compared to the 6C CCCV strategy. This is mainly attributed to their well-constrained temperature and side reaction overpotential in the charging process. In contrast, the 6C CCCV protocol is observed to incur fast degradation even at the early life stage due to the unawareness of degradation protection. Given a similar charging

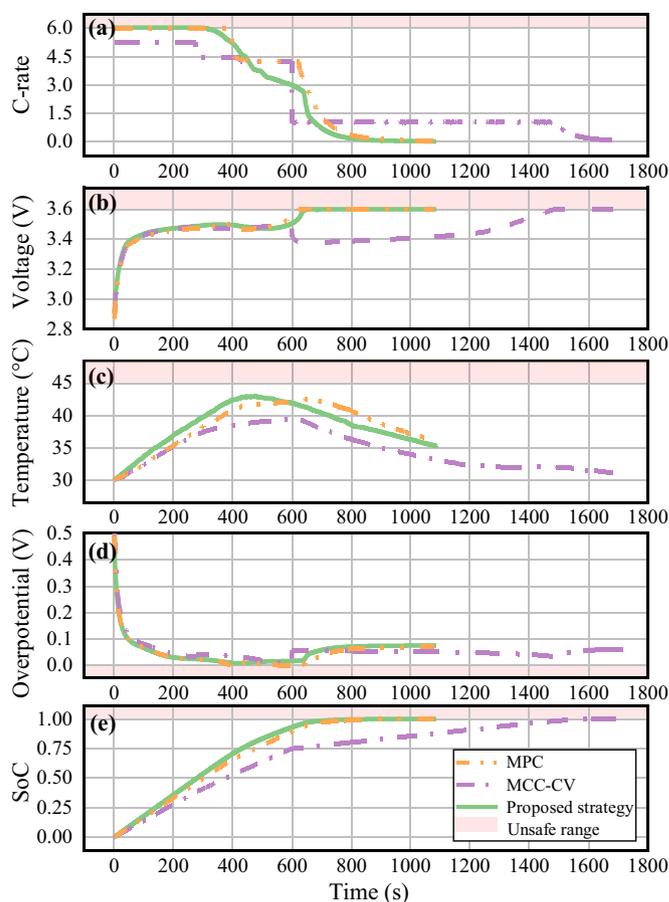


Fig. 9. Comparison of the proposed strategy with the other advanced strategies: (a) current, (b) terminal voltage, (c) core temperature, (d) side reaction overpotential, (e) SoC.

Table 7

Comparison of the consumed charge time for the different strategies.

Strategy	4C CCCV	6C CCCV	MPC	MCC–CV	Proposed strategy
To 80% SoC (s)	742	489	520	797	471
To 90% SoC (s)	834	550	607	1173	569
To 100% SoC (s)	1300	901	1078	1723	1087

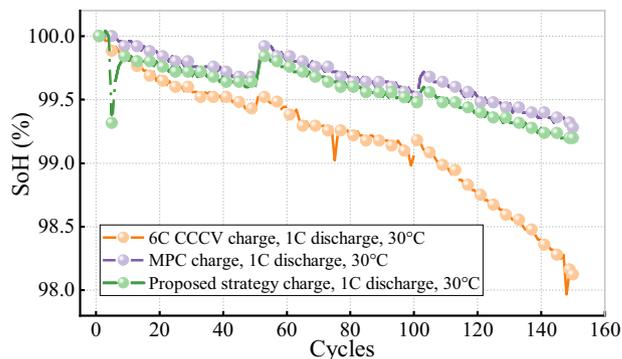


Fig. 10. Results of the long-term cycling experiments for three strategies.

speed, the cycle life has been extended by about 75% using the proposed strategy, compared to the case of using 6C CCCV protocol. It is also meaningful to evaluate the cycle life of batteries by using different

Table 8
Results of predicted cycle life for the different strategies.

Strategy	6C CCCV	MPC	Proposed strategy
cycle life	599	1044	1010

charging strategies. To this end, the data-driven approaches can be used to predict the remaining useful life with only early-life degradation data [36,37]. Herein, the total cycle life of batteries are predicted using the time-series-based data-driven approach, leveraging the capacity fading data from the first 100 cycles. The predicted results are shown in Table 8. Consistent with the previously shown results, the proposed strategy shows a similar cycle life with MPC and outperforms 6C CCCV strategy significantly. In summary, the experimental results validate the superiority of proposed strategy regarding the optimal balance between the charging rapidity and anti-degradation performance.

It is worth noting that the ROM and charging strategy need to be updated along with the battery aging according to the aforementioned battery-to-cloud framework. From the perspective of cycling characteristics, however, the degradation of LIB happens in a long timescale while using the proposed strategy, which implies that the changes in ROM parameters are not significant over a long period of time. Specifically, comparing the modeling errors for fresh cell with and the aged cell after 150 cycles under initial parameters, the results reveal that the mean absolute errors of the voltage and temperature only slightly increase by 0.005 V and 0.02 °C. Such small model error is almost harmless to the proposed strategy. Therefore, the strategy can work well without frequent re-calibration. Even in the case of long-term operation where the significant degradation happens, the periodical update of model and charging strategy is easily feasible with the battery-to-cloud architecture. Specifically, the time-consuming model re-calibration and strategy re-training will be performed periodically in the cloud server, while the re-trained charging strategy will be downloaded and implemented in the real-time controller for practical charging control.

4.4. Comparison with state-of-the-art DRL methods

Two popular DRL algorithms, i.e., deep Q-network (DQN) and deep deterministic policy gradient (DDPG), are used as benchmarks to further justify the performances of the proposed strategy. For DQN-based strategy, the action space is dispersed into 20 intervals to balance the precision and convergence. The training sets of three strategies are originated from the same environment for a fair comparison. Specially, each strategy is trained five times under different initial networks (by changing random seeds) to reduce the potential contingency of training.

The training results are given comparatively in Fig. 11, where the solid line and shadow represents the average value and boundaries of

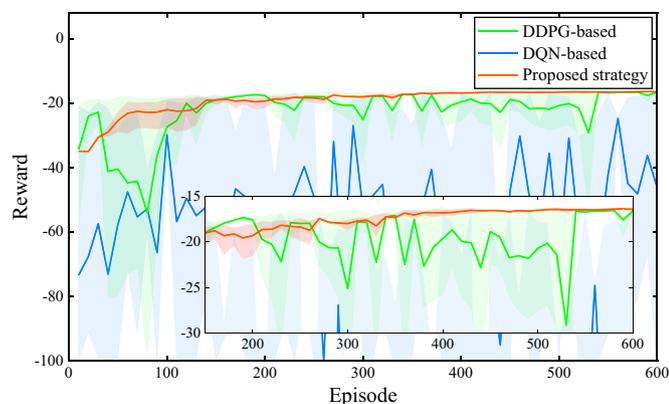


Fig. 11. Comparison of training results on the proposed strategy with state-of-the-art DRL-based benchmarks.

reward for the five trainings. Apparently, the three DRL algorithms show remarkable differences in the convergence properties. It is observed that the reward increases rapidly and stabilizes along with the growing episodes for the DDPG and the proposed strategy, implying the potential convergence of two strategies. In contrast, the reward given by the DQN strategy shows successive and large fluctuations over the entire training process, which suggests that the discrete DQN can hardly address the continuous control problems. It is also shown that the converging rate and the maximum reward of the DDPG are both notably less than the proposed strategy. Moreover, under different initialization conditions, the reward of DDPG exhibits a wider fluctuation range than that of the proposed strategy. This can be explained by the inherent severe parameter sensitivity and over-estimation problem of DDPG. By comparison, the proposed strategy reduces the dependence of action on the Q-function by introducing the maximum entropy, which alleviates the parameter sensitivity effectively.

To verify the above conjecture, the episode with the best convergence under each strategy is further selected for comparison, and the results are depicted in Fig. 12. Apparently, the DQN strategy fails to find the optimal charging trajectory, since the early stage of charging has been determined with a small charging current, and the overall charging is much longer than the other two strategies. The charging current appears some inherent step jumps due to the dispersion of actions. Explicitly, the DDPG exhibits improved charging performance compared with the DQN. However, the DDPG strategy incurs an unsafe voltage overshoot. This exactly explains the lower maximum reward of DDPG after convergence. By comparison, the proposed strategy has shown expected optimality and stability. The charging speed has been promoted greatly while the physical limitations of LIB have not been

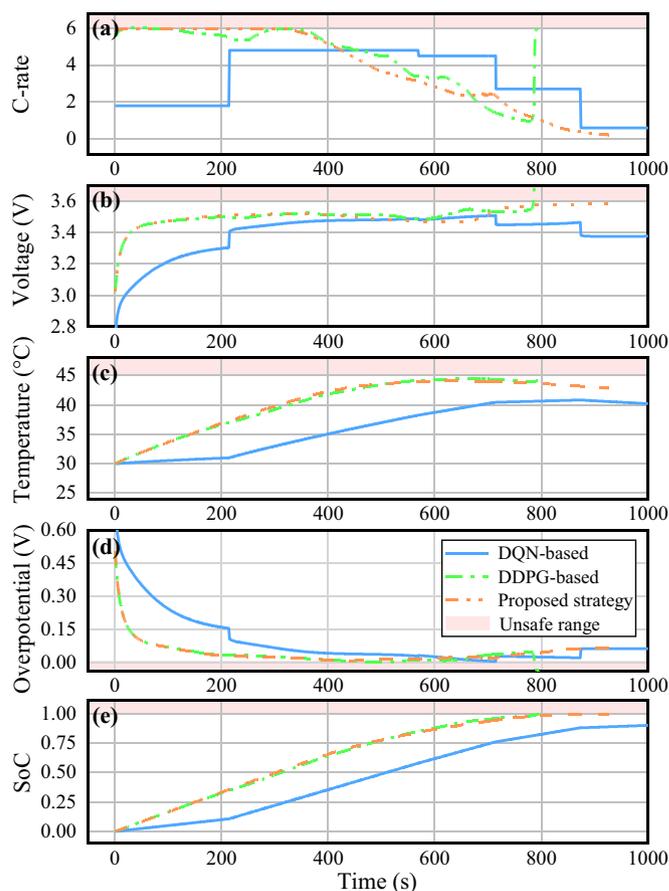


Fig. 12. Comparison of the proposed strategy with the other DRL-based strategies: (a) current, (b) terminal voltage, (c) core temperature, (d) side reaction overpotential, (e) SoC.

violated.

4.5. Discussion

In theory, the complicated mechanisms involved in LIB cannot be covered by any known battery models to date, and the model in this work is no exception. However, the proposed machine learning-based charging strategy is still meaningful taking account of the following points.

- (1) Albeit not considering all the essential mechanism thoroughly, the proposed model has been well validated for the accurate prediction of terminal voltage, surface temperature and internal temperature of LIB. Upon modeling these parameters/states precisely, the proposed AI-based charging strategy is reliable for constraining the corresponding abusive conditions of over-voltage and over-heat. This is a major contribution of the proposed strategy.
- (2) The side-reaction overpotential has been calculated with well-recognized governing equations widely used over years. Although not measured directly, the modeling of side-reaction overpotential is considered to be approaching the ground truth. At least within the available knowledge, the developed model and its associated overpotential-constraining strategy can be insightful for lithium plating prevention during LIB fast charging. More validations and modeling techniques will be studied to enhance this part in the future work.
- (3) Although the developed model and charging strategy can only consider very limited degradation mechanisms, the comparative long-term cycling experiments explicitly validates the superiority of the proposed strategy over the state of the art. The proposed machine learning-based approach validates to keep an equivalent charging speed while extends the battery life remarkably compared to the commonly-used charging method.
- (4) The proposed charging strategy is highly compatible to any category of battery models. That means the charging performance can be easily upgraded once more-advanced battery models can be developed. The investigation of refined battery models capturing more degradation mechanisms has always been a hot research topic over years. Although the modeling is not the major endeavor of this paper, the future emergence of advanced models can be easily implanted within the proposed framework of fast charging. This is another contribution justifying the value of the proposed method.

5. Conclusions

The paper proposes a machine learning-based fast charging strategy for the LIB with safety and health consciousness by perceiving and regulating the key microscopic states. The strategy incorporates an ROM for microscopic state perceiving and the state-of-the-art SAC-DRL algorithm for policy training. The proposed strategy is performed within a battery-to-cloud framework, where the ROM-enabled policy training is implanted in the cloud, while the trained low-cost strategy is performed in the real-time controller to avoid the risk of a heavy computing demand. The major conclusions are drawn as follows.

- (1) Attributed to the ROM-enabled microscopic state perceiving, the proposed strategy can accelerate the charging utmostly, while the core temperature and side reaction overpotential are confined smartly with a expected region, so that the unfavorable over-temperature and lithium deposition can be mitigated during the fast charging.
- (2) Given a similar charging speed, the proposed machine learning-based approach extends the battery cycle life by about 75% compared to the commonly-used rule-based charging protocol,

which highlights the benefit of health awareness and active anti-aging control.

- (3) Compared with the state-of-the-art rule-based and model-based strategies, the proposed strategy performs equivalently with the MPC one in terms of charging optimality and significantly outperforms MCC—CV one. However, the proposed strategy is about two orders of magnitude lower than MPC one in time complexity, providing a better potential for real-time utilization.
- (4) The proposed strategy is highly adaptive to the uncertain environment. Encouraging charging performances have been obtained under different ambient temperatures and initial SoCs. Compared to the state-of-the-art DRL algorithms, the SAC-DRL strategy exhibits appealing performance regarding the converging property, stability, and optimality.

CRediT authorship contribution statement

Zhongbao Wei: Conceptualization, Methodology, Investigation, Resources, Writing – original draft, Supervision, Funding acquisition. **Xiaofeng Yang:** Methodology, Validation, Formal analysis, Investigation, Writing – review & editing. **Yang Li:** Methodology, Formal analysis, Investigation, Writing – review & editing. **Hongwen He:** Formal analysis, Investigation, Writing – review & editing. **Weihan Li:** Writing – review & editing. **Dirk Uwe Sauer:** Conceptualization, Resources, Writing – review & editing.

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.

Data availability

Data will be made available on request.

Acknowledgement

This work is supported by the National Key R&D Program of China under Grant 2022YFB2405700, and the National Natural Science Foundation of China under Grant U22A20227.

References

- [1] R. Schmich, R. Wagner, G. Hörpel, T. Placke, M. Winter, Performance and cost of materials for lithium-based rechargeable automotive batteries, *Nature Energy* 3 (2018) 267–278.
- [2] L. Xu, X. Lin, Y. Xie, X. Hu, Enabling high-fidelity electrochemical P2D modeling of lithium-ion batteries via fast and non-destructive parameter identification, *Energy Storage Mater.* 45 (2022) 952–968.
- [3] D.A. nseán, M. Dubarry, A. Devie, B.Y. Liaw, V.M. García, J.C. Viera, et al., Fast charging technique for high power LiFePO4 batteries: a mechanistic analysis of aging, *J. Power Sources* 321 (2016) 201–209.
- [4] A. Tomaszewska, Z. Chu, X. Feng, S. O’Kane, X. Liu, J. Chen, et al., Lithium-ion battery fast charging: a review, *eTransportation* 1 (2019).
- [5] S. Ahmed, I. Bloom, A.N. Jansen, T. Tanim, E.J. Dufek, A. Pesaran, et al., Enabling fast charging – a battery technology gap assessment, *J. Power Sources* 367 (2017) 250–262.
- [6] S.S. Zhang, K. Xu, T.R. Jow, Study of the charging process of a LiCoO₂-based Li-ion battery, *J. Power Sources* 160 (2006) 1349–1354.
- [7] T.T. Vo, X. Chen, W. Shen, A. Kapoor, New charging strategy for lithium-ion batteries based on the integration of Taguchi method and state of charge estimation, *J. Power Sources* 273 (2015) 413–422.
- [8] Y.H. Liu, C.H. Hsieh, Y.F. Luo, Search for an optimal five-step charging pattern for li-ion batteries using consecutive orthogonal arrays, *IEEE Trans. Energy Conver.* 26 (2011) 654–661.
- [9] P.M. Attia, A. Grover, N. Jin, K.A. Severson, T.M. Markov, Y.H. Liao, et al., Closed-loop optimization of fast-charging protocols for batteries with machine learning, *Nature* 578 (2020) 397–402.
- [10] P. Notten, J. Veld, J. Beek, Boostcharging Li-ion batteries: a challenging new charging concept, *J. Power Sources* 145 (2005) 89–94.

- [11] L. Chen, S. Wu, D. Shieh, T. Chen, Sinusoidal-ripple-current charging strategy and optimal charging frequency study for Li-Ion batteries, *IEEE Trans. Ind. Electron.* 60 (2013) 88–97.
- [12] F. Savoye, P. Venet, M. Millet, J. Groot, Impact of periodic current pulses on Li-Ion battery performance, *IEEE Trans. Ind. Electron.* 59 (2012) 3481–3488.
- [13] K. Liu, K. Li, H. Ma, J. Zhang, Q. Peng, Multi-objective optimization of charging patterns for lithium-ion battery management, *Energy Convers. Manag.* 159 (2018) 151–162.
- [14] K. Liu, X. Hu, Z. Yang, Y. Xie, S. Feng, Lithium-ion battery charging management considering economic costs of electrical energy loss and battery degradation, *Energy Convers. Manag.* 195 (2019) 167–179.
- [15] M. Xu, R. Wang, P. Zhao, X. Wang, Fast charging optimization for lithium-ion batteries based on dynamic programming algorithm and electrochemical-thermal-capacity fade coupled model, *J. Power Sources* (2019) 438.
- [16] Y. Yin, Y. Hu, S.-Y. Choe, H. Cho, W.T. Joe, New fast charging method of lithium-ion batteries based on a reduced order electrochemical model considering side reaction, *J. Power Sources* 423 (2019) 367–379.
- [17] P. Nambisan, P. Saha, M. Khanra, Real-time optimal fast charging of Li-ion batteries with varying temperature and charging behaviour constraints, *J. Energy Storage* 41 (2021), 102918.
- [18] C. Zou, X. Hu, Z. Wei, X. Tang, Electrothermal dynamics-conscious lithium-ion battery cell-level charging management via state-monitored predictive control, *Energy* 141 (2017) 250–259.
- [19] Z. Chu, X. Feng, L. Lu, J. Li, X. Han, M. Ouyang, Non-destructive fast charging algorithm of lithium-ion batteries based on the control-oriented electrochemical model, *Appl. Energy* 204 (2017) 1240–1250.
- [20] C. Zou, X. Hu, Z. Wei, T. Wik, E. Bo, Electrochemical estimation and control for lithium-ion battery health-aware fast charging, *IEEE Trans. Ind. Electron.* 65 (2018) 6635–6645.
- [21] Y. Gao, X. Zhang, B. Guo, C. Zhu, J. Wiedemann, L. Wang, et al., Health-aware multiobjective optimal charging strategy with coupled electrochemical-thermal-aging model for lithium-ion battery, *IEEE Trans. Ind. Inf.* 16 (2020) 3417–3429.
- [22] W. Li, I. Demir, D. Cao, D. Jöst, F. Ringbeck, M. Junker, et al., Data-driven systematic parameter identification of an electrochemical model for lithium-ion batteries with artificial intelligence, *Energy Storage Mater.* 44 (2022) 557–570.
- [23] V.R. Subramanian, V.D. Diwakar, D. Tapriyal, Efficient Macro-Micro scale coupled modeling of batteries, *J. Electrochem. Soc.* (2005) 152.
- [24] W. Luo, C. Lyu, L. Wang, L. Zhang, An approximate solution for electrolyte concentration distribution in physics-based lithium-ion cell models, *Microelectron. Reliab.* 53 (2013) 797–804.
- [25] E. Prada, D. Di Domenico, Y. Creff, J. Bernard, V. Sauvant-Moynot, F. Huet, Simplified electrochemical and thermal model of LiFePO₄-graphite Li-Ion batteries for fast charge applications, *J. Electrochem. Soc.* 159 (2012) A1508–A1A19.
- [26] C. Forgez, D. Vinh Do, G. Friedrich, M. Morcrette, C. Delacourt, Thermal modeling of a cylindrical LiFePO₄/graphite lithium-ion battery, *J. Power Sources* 195 (2010) 2961–2968.
- [27] W. Mei, L. Zhang, J. Sun, Q. Wang, Experimental and numerical methods to investigate the overcharge caused lithium plating for lithium ion battery, *Energy Storage Mater.* 32 (2020) 91–104.
- [28] C. Edouard, M. Petit, C. Forgez, J. Bernard, Revel R. Parameter sensitivity analysis of a simplified electrochemical and thermal model for Li-ion batteries aging, *J. Power Sources* 325 (2016) 482–494.
- [29] M. Safari, C. Delacourt, Modeling of a commercial Graphite/LiFePO₄ Cell, *J. Electrochem. Soc.* (2011) 158.
- [30] Z. Wei, J. Hu, H. He, Y. Yu, J. Marco, Embedded distributed temperature sensing enabled multi-state joint observation of smart lithium-ion battery, *IEEE Trans. Ind. Electron.* 70 (2023) 555–565.
- [31] Z. Wei, P. Li, W. Cao, H. Chen, W. Wang, Y. Yue, H. He, Hongwen He. Machine learning-based hybrid thermal modeling and diagnostic for lithium-ion battery enabled by embedded sensing, *Appl. Therm. Eng.* 216 (2022), 119059.
- [32] M.S.S. Malik, G. Li, Z. Chen, An optimal charging algorithm to minimise solid electrolyte interface layer in lithium-ion battery, *J. Power Sources* (2021) 482.
- [33] T.P. Lillicrap, J.J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, et al., Continuous control with deep reinforcement learning, *Comput. Sci.* (2015).
- [34] W. Li, M. Rentemeister, J. Badedá, D. Jöst, D. Schulte, D.U. Sauer, Digital twin for battery systems: cloud battery management system with online state-of-charge and state-of-health estimation, *J. Energy Storage* 30 (2020), 101557.
- [35] P. Lyu, X. Liu, J. Qu, J. Zhao, Y. Huo, Z. Qu, et al., Recent advances of thermal safety of lithium ion battery for energy storage, *Energy Storage Mater.* 31 (2020) 195–220.
- [36] X. Liu, X.-Q. Zhang, X. Chen, G.-L. Zhu, C. Yan, J.-Q. Huang, et al., A generalizable, data-driven online approach to forecast capacity degradation trajectory of lithium batteries, *J. Energy Chem.* 68 (2022) 548–555.
- [37] J. Tian, R. Xiong, W. Shen, J. Lu, X.-G. Yang, Deep neural network battery charging curve prediction using 30 points collected in 10min, *Joule* 5 (2021) 1521–1534.