



Lifelong Reinforcement Learning for Health-Aware Fast Charging of Lithium-ion Batteries

Downloaded from: <https://research.chalmers.se>, 2025-11-09 18:54 UTC

Citation for the original published paper (version of record):

Yuan, M., Zou, C. (2025). Lifelong Reinforcement Learning for Health-Aware Fast Charging of Lithium-ion Batteries. IEEE Transactions on Transportation Electrification, 1.
<http://dx.doi.org/10.1109/TTE.2025.3625421>

N.B. When citing this work, cite the original published paper.

© 2025 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, or reuse of any copyrighted component of this work in other works.

Lifelong Reinforcement Learning for Health-Aware Fast Charging of Lithium-ion Batteries

Meng Yuan, *Member, IEEE*, Changfu Zou, *Senior Member, IEEE*

Abstract—Fast charging of lithium-ion batteries remains a critical bottleneck for widespread adoption of electric vehicles and stationary energy storage systems, as improperly designed fast charging can accelerate battery degradation and shorten lifespan. In this work, we address this challenge by proposing a health-aware fast charging strategy that explicitly balances charging speed and battery longevity across the entire service life. The key innovation lies in establishing a mapping between side-reaction overpotential and the state of health (SoH) of battery, which is then used to constrain the terminal charging voltage in a twin delayed deep deterministic policy gradient (TD3) framework. By incorporating this SoH-dependent voltage constraint, our designed deep learning method mitigates side reactions and effectively extends battery life. To validate the proposed approach, a high-fidelity single particle model with electrolyte is implemented in the widely adopted PyBaMM simulation platform, capturing degradation phenomena at realistic scales. Comparative life-cycle simulations against conventional CC-CV, its variants, and constant current–constant overpotential methods show that the TD3-based controller reduces overall degradation while maintaining competitively fast charge times. These results demonstrate the practical viability of deep reinforcement learning for advanced battery management systems and pave the way for future explorations of health-aware, performance-optimized charging strategies.

Index Terms—Lithium-ion battery, fast charging, reinforcement learning, battery degradation.

I. INTRODUCTION

LITHIUM-ION batteries (LIBs) have emerged as the cornerstone of modern energy storage systems, powering applications ranging from portable electronics to electric vehicles (EVs) and grid-scale renewable energy integration [1]. Despite their dominance, EVs continue to face two persistent challenges, limited driving range and prolonged charging durations, which remain critical barriers to their widespread adoption and the push for energy-efficient electrification [2], [3]. The energy density, charge and discharge kinetics of LIBs are primarily determined by their electrochemical composition, including cathode and anode materials and electrolytes, as well as their structural design, such as particle morphology and electrode architecture [4]. While advancements in battery

chemistry continue to push theoretical limits of battery, sophisticated battery management systems (BMS) have shown promise in optimizing charging protocols to achieve fast charging without exceeding intrinsic material constraints [5]. Though effective in reducing charging time, unconstrained high-current charging accelerates capacity fade through mechanisms such as lithium plating and solid-electrolyte interphase (SEI) growth [6]. To address this dilemma, advanced control algorithms capable of dynamically balancing charging speed and longevity, through multi-physics-aware current modulation and degradation-predictive interventions, are urgently required to unlock the full potential of LIBs [7]. In this context, we approach fast charging from a lifelong, health-aware perspective. Unlike prior studies that emphasize only short-term speed or implicitly account for degradation, we explicitly incorporate long-term battery health into the charging problem. This high-level perspective highlights the novelty of our work and motivates the methodological development described in the following sections.

Accurate battery modeling is essential for designing health-aware fast charging control algorithms. One widely used approach is the equivalent circuit model (ECM), which simulates voltage dynamics using simplified resistor-capacitor networks and offers computational efficiency ideal for real-time embedded systems [8], [9]. However, without modeling the first principles underlying intercalation reactions and the diffusion process within battery cells, ECMs often struggle to capture locally distributed behavior, particularly those related to safety and aging. In contrast, electrochemical models, such as the Doyle-Fuller-Newman (DFN) model, explicitly describe ion transport and reaction kinetics across electrode microstructures, enabling physics-based predictions of degradation modes [10]. Yet, the significant computational burden of solving the coupled partial differential equations (PDEs) inherent in the DFN models limits their practical application in advanced control methods that also demand high computational power. To address these challenges, the single particle model (SPM) has emerged as a pragmatic simplification. By assuming a uniform electrolyte concentration, the SPM reduces PDEs to a set of ordinary differential equations (ODEs) while still capturing essential electrochemical behavior [11]. Recent work has extended the SPM to incorporate various capacity fade mechanisms and integrated it into a model predictive control (MPC) framework to optimize charging profiles and minimize intra-cycle capacity fade [12]. However, the conventional SPM neglects electrolyte dynamics and leads to significant voltage prediction errors at high C-rates, which has motivated the development of extensions such as the single

This work was supported by the European Union's Horizon Europe programme under the Marie Skłodowska-Curie Actions Postdoctoral Fellowships (Grant No. 101110832), and the Swedish Research Council (Grant No. 2023-04314). The mobility of C. Zou was funded by the Swedish Foundation for International Cooperation in Research and Higher Education (STINT; Grant No. MG2020-9107), and the Science and Technology Innovation program of Hunan Province. (Corresponding author: Changfu Zou.)

Meng Yuan and Changfu Zou are with the Department of Electrical Engineering, Chalmers University of Technology, 412 96 Gothenburg, Sweden. (Emails: meng.yuan@ieee.org; changfu.zou@chalmers.se)

particle model with electrolyte (SPMe) for improved accuracy [13], [14].

Conventional lithium-ion battery charging typically employs the constant current–constant voltage (CC-CV) method, where the battery is first charged at a fixed current until a set voltage is reached, and then held at constant voltage while the current gradually decreases. Although fast charging can be achieved by increasing the current or target voltage, such measures tend to accelerate battery degradation phenomena such as lithium plating, thereby reducing battery life [15]. To solve this issue, researchers turned to health-aware fast charging strategies that integrate battery degradation models with advanced control techniques. In [16], a side-reaction overpotential based PID controller is designed to mitigate the risks of lithium plating while increasing the speed of charging. In [17], Yang et al. proposed a constant-overpotential based fast charging strategy for Li-ion batteries and show that this constant-overpotential control outperforms the traditional CC-CV protocol in both charging speed and lithium-plating suppression. Similarly, Lu et al. developed a MPC framework that incorporates real-time lithium plating detection and adaptive parameter updates to optimize the balance between fast charging and long-term battery health [18]. While model-based strategies like MPC are powerful for constrained optimization, their application to lifelong fast charging presents significant challenges. Specifically, the performance of MPC is critically dependent on model fidelity, requiring robust online parameter estimation to track the changing dynamics of the battery as it ages. As demonstrated in [19], an optimal control strategy based on a non-adaptive model can paradoxically accelerate aging compared to a standard CC-CV protocol. The complexity of integrating a high-fidelity aging model, an adaptive estimation scheme, and real-time optimization is a considerable barrier to practical deployment of MPC in this application.

Moreover, the side-reaction overpotential is hard to be measured in practical application, which hinders the implementation of overpotential based method. This challenge has inspired the development of various approaches including machine learning-based methods, to estimate lithium plating [20]. Beyond LIBs, accurate operando detection of lithium plating during high-rate charging has also been demonstrated in lithium-ion capacitors. Although the chemistry differs, these results underscore the importance of plating-aware diagnostics at extreme C-rates and motivate detection-informed or constraint-based charging strategies [21].

To address the challenges of model mismatch and parameter drift as batteries age, learning-based methods have been introduced in the field of battery charging. In [22], a deep reinforcement learning (RL) based approach has been designed for the fast charging of battery, and two cases considering if the side-reaction overpotential is measurable are discussed in the design of the controller. While the method shows promise for rapid charging, it models battery aging solely as an increase in film resistance, and the controller design does not integrate long-term battery degradation considerations. In [23], Wei et al. propose a deep RL-based strategy for the fast charging of lithium-ion batteries, specifically targeting thermal safety and health-conscious charging. The multiphysics-related

constraints are implicitly incorporated into the reward design of the RL. However, the agent is trained on a per-episode basis, with the battery model initialized at the beginning of each epoch, overlooking the long-term aging dynamics that recent work has identified as critical for developing lifelong strategies [24], [25]. In [26], an adaptive model-based RL strategy leverages Gaussian processes to capture the battery environment and enforce operational constraints during charging, but its primary focus is on minimizing charging time rather than extending battery life.

Based on the discussion above, our work addresses the challenge in fast charging by taking battery degradation into explicit account throughout the entire lifespan. We employ a twin delayed deep deterministic policy gradient (TD3) method for health-aware fast charging. To mitigate battery degradation, we first establish a mapping between the side-reaction overpotential and the charge cut-off voltage and then integrate it as health-related constraint into the training of our RL agent. The SPM coupling degradation is employed as the battery environment to capture electrochemical behavior. The contributions of this work are summarized as follows:

- Unlike previous RL-based approaches that focus solely on minimizing charging time with implicit health constraints, or that train the RL agent by resetting the battery to a fresh state after each training episode, this work introduces the first explicit formulation of a lifelong battery fast charging problem. This is achieved through a new training procedure where the RL agent was trained on a battery model that ages continuously, enabling a strategy that truly reduces charging duration while extending battery longevity over its entire service life.
- A mapping between the charge cut-off voltage and SoH was established using a constant current (CC)–constant overpotential (COP) approach. This mapping leverages the intrinsic relationship between the side-reaction overpotential and battery degradation, and can be integrated into both CCCV-based charging and learning-based advanced charging, allowing directly incorporation of battery health into charging control.
- By integrating the obtained mapping between the voltage and SoH, a TD3-based charging strategy is designed to explicitly optimize the fast charging process.
- The proposed TD3-based strategy is trained and tested through comprehensive life-cycle simulations using the battery simulator Python Battery Mathematical Modeling (PyBaMM) [27] with a SPM coupling degradation model. Superior performance is demonstrated over CC-CV charging and its variants.

The remainder of this paper is organized as follows. Section II defines the health-aware fast charging problem. Section III presents the TD3-based reinforcement learning approach. Section IV reports modeling details, experimental setup, and charging results, including analysis of aging effects and lithium plating. Finally, Section V concludes the paper.

II. HEALTH-AWARE FAST CHARGING PROBLEM

A battery is considered to have reached its end of life when its SoH falls to a specified threshold, such as 75-80%

for EV applications. Within this health-aware fast charging framework, we aim to minimize the charging time needed to reach the desired SoC while maximizing the number of equivalent full cycles (EFCs) of battery until the SOH reaches 80%. The cost function for this problem is defined as:

$$J(t) = w_1 C_{\text{SoC}}(t) + w_2 C_{\text{side}}(t), \quad (1)$$

where w_1 and w_2 are weighting factors. The first part of the cost is designed for SoC setpoint tracking, aiming to minimize charging duration. The cost function is defined as:

$$C_{\text{SoC}}(t) = |\text{SoC}^* - \text{SoC}(t)|, \quad (2)$$

where SoC^* denotes the target state of charge, and SoC_t represents the current state of charge. Considering that lithium plating is the main degradation mechanism here, the degradation-related cost is defined as:

$$C_{\text{side}}(t) = |\eta_{\text{side}}(t) - \eta_{\text{min}}|, \quad (3)$$

where η_{min} is a predefined threshold overpotential, below which battery degradation becomes excessively rapid. We can interpret this threshold as follows: when the side-reaction overpotential η_{side} falls below this value, further reducing it to achieve faster charging is no longer beneficial, as the resulting acceleration in battery degradation outweighs any additional gains in charging speed.

The constraint of this problem involves that the input current is limited by the battery charger as:

$$0 \leq I(t) \leq I_{\text{max}}. \quad (4)$$

Then, the health-aware fast charging problem is formulated as

$$\begin{aligned} \min_{u(t)} \quad & \int_0^{t_e} J(t) dt \\ \text{s.t.} \quad & \text{Battery dynamics: } \dot{x}(t) = f(x(t), u(t)), \\ & \text{Input constraints: } (4), \\ & \text{SoC}(t_f) = \text{SoC}^*, \quad \forall \text{ cycles}, \\ & \text{SoH}(t_e) = \text{SoH}_{\text{end}}, \end{aligned} \quad (5)$$

where x denotes the state of battery dynamics, u is the charging current serving as the system input, t_f is the charging completion time for each cycle, which can vary over the lifespan of battery, t_e represents the end-of-life time of the battery, and SoH_{end} is the specified battery SoH at end-of-life, set to 80% in this study.

This optimization problem is formulated to minimize the overall cost function over the battery lifetime from initial time 0 to its end-of-life time t_e . Each charging cycle within this period terminates at time t_f when the SoC reaches the target SoC^* .

A. Constant current constant overpotential based control

To solve the fast charging problem as shown in (5), one intuitive idea is to design the controller based on side-reaction overpotential feedback. This constant current-constant overpotential (CC-COP) method mirrors the traditional CC-CV

approach by initiating with a constant current phase, followed by dynamically regulating the side-reaction potential to a small positive value using real-time feedback. This method has been shown to effectively reduce the risk of plating by maintaining the side-reaction overpotential at a safe, constant level after the initial current phase [17], [20].

However, the CC-COP control in fast charging faces two significant limitations: Firstly, measuring electrode potentials directly in commercial batteries is challenging as it requires advanced sensors typically used only in laboratory settings, making it impractical for widespread application. Secondly, determining the appropriate side-reaction overpotential threshold is not straightforward. While setting a high threshold can help reduce lithium plating, it may result in overly conservative charging corresponding to very long charging times. The need to finely tune the side-reaction overpotential threshold adds complexity to real-world implementation, further complicating the adoption of CC-COP control strategies in commercial battery charging systems.

III. REINFORCEMENT LEARNING BASED FAST CHARGING ALGORITHM

To overcome the aforementioned limitations of CC-COP control, a data-driven alternative based on RL is adopted. Unlike model-based strategies that depend on the development and calibration of accurate physical models, RL learns optimal charging policies directly through interactions with the environment, using observable signals and reward feedback. This paradigm enables the development of adaptive charging strategies that can incorporate battery health indicators without requiring explicit physical modeling.

Fast charging control requires a continuous, bounded charging current and is trained in a high-fidelity environment where samples are costly. For safe deployment in BMS, we also prefer low-variance, smooth actions and stable long-horizon learning. These requirements favor an off-policy, sample-efficient deterministic actor-critic method with mechanisms that curb Q-value overestimation. We therefore adopt TD3 [28], a stabilized successor of DDPG that combines twin critics [29], target-policy smoothing, and delayed actor updates to improve stability and precision in continuous control.

In this work, we design a controller that builds upon the insights of CC-COP and implement it within the above TD3 framework to solve the health-aware fast charging problem defined in (5). Fig. 1 illustrates the overall system structure, which comprises three main components.

The first component is the battery environment, which may represent either a real battery system or a high-fidelity lithium-ion battery model that captures degradation dynamics. Since the side-reaction overpotential is challenging to be measured in practical applications, we aim to acquire a correlation between the cut-off voltage and SoH based on the CC-COP controller, as detailed in Section II-A. The maximum terminal voltage of the battery cell, $V_{\text{cut-off}}$, during each charging process that consistently aims for the SoC setpoint, SoC^* is recorded during charging. This voltage identification continues until battery SoH falls below 80%, signaling the end of its lifespan.

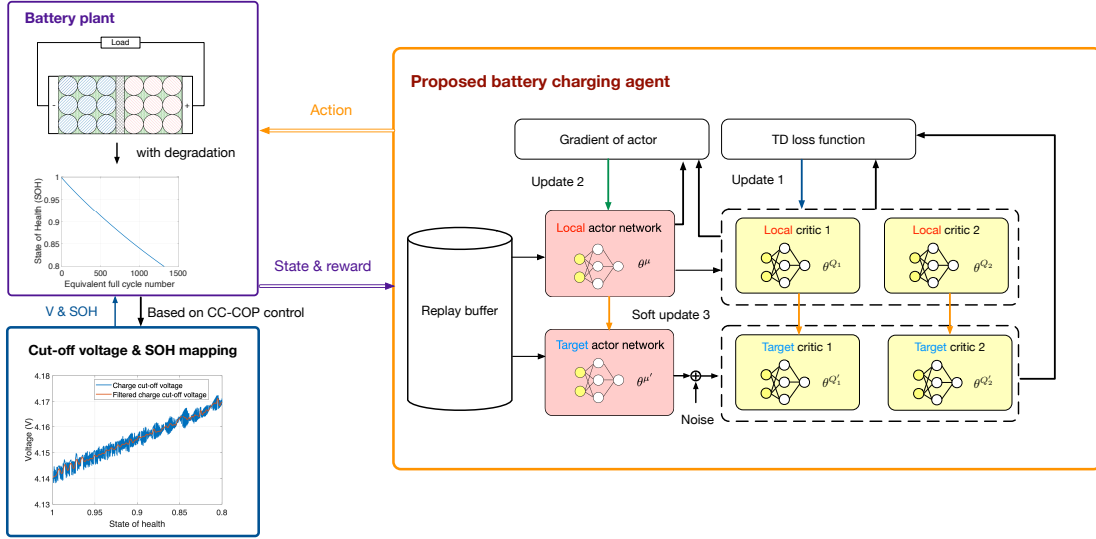


Fig. 1: Structure of the proposed method.

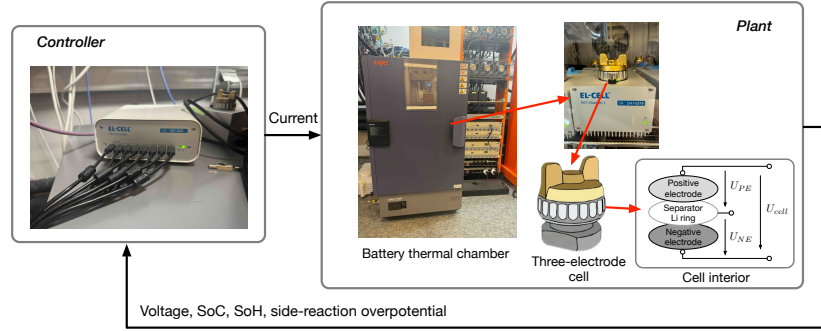


Fig. 2: Experimental setup for CC-COP battery cycling.

This step can be experimentally performed using the setup shown in Fig. 2, where a PAT-Cell system from EL-CELL is employed for testing three-electrode cells, or via a high-fidelity simulation. The resulting correlation between charge cut-off voltage and SoH provides important battery health information to the environment, enabling the generation of reward signals for the battery charging agent. The third and core component of our system is the TD3-based charging agent, which receives state and reward information from the environment and computes optimal charging actions accordingly. The detailed formulation and implementation of this agent is presented in the following sections.

For practical implementation of the proposed method, building the $V_{\text{cutoff}}\text{--SoH}$ map using the CC-COP strategy requires a direct measurement of the side-reaction overpotential. Although such a measurement may not be available in some commercial two-electrode cells, the challenge can be addressed by performing a one-time, offline characterization using a representative three-electrode laboratory cell. For this process, electrode materials can be harvested from commercial cells inside a glovebox and subsequently reassembled. This enables the precise, direct measurement of η_{side} , which is otherwise inaccessible. By cycling this laboratory cell, we obtain a high-fidelity $V_{\text{cutoff}}\text{--SoH}$ map that serves as the ground truth

for deployment. While this characterization must be repeated for each new battery type, we identify machine learning, particularly transfer learning, as a promising direction for future work to accelerate map generation and reduce the experimental workload.

A. TD3 method overview

To address the charging problem specified in (5), the TD3 algorithm, an enhanced version of the DDPG framework, aims to rectify the Q-value overestimation issue inherent in DDPG. It introduces targeted improvements that boost the learning stability and accuracy. The architecture of the TD3 algorithm is illustrated in Fig. 3. Within the Actor-Critic framework, TD3 incorporates two types of networks: the critic network $Q(s, a|\theta^Q)$ with parameter θ^Q and the actor network $\mu(s|\theta^\mu)$ with parameter θ^μ . To mitigate the overestimation of the Q-value, TD3 employs a technique that calculates the target Q-value by selecting the lesser value from the two target critic networks, effectively reducing potential overestimation:

$$y_t(r, s') = r(s, a) + \gamma \min_{i=1,2} Q'_i(s', a'|\theta^{Q'_i}), \quad (6)$$

where y_t is the target Q-value for given state s and action a ; a' is the next action chosen by the policy based on next state s' ; γ is the discounting factor; $\theta^{Q'_i}$ is the parameter of

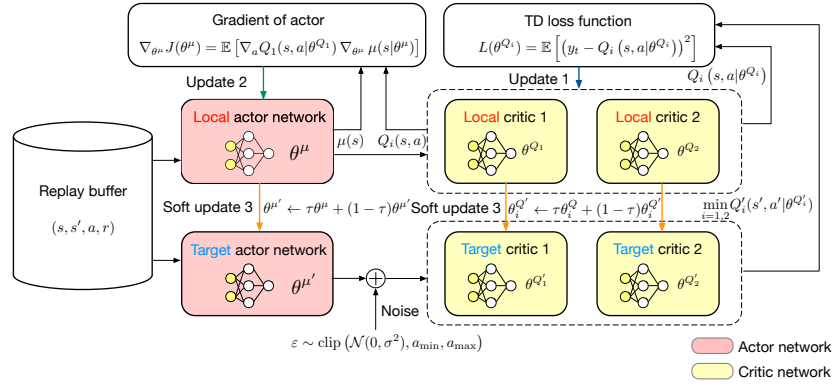


Fig. 3: Structure of the deep learning algorithm.

target critic. The critic network updates its parameters based on gradient descent applied to the loss function as follows:

$$L(\theta^{Q_i}) = \mathbb{E} \left[(y_t - Q_i(s, a | \theta^{Q_i}))^2 \right], \quad (7)$$

$$\nabla_{\theta^{Q_i}} L(\theta^{Q_i}) = \mathbb{E} \left[(y_t - Q_i(s, a | \theta^{Q_i})) \nabla_{\theta^{Q_i}} Q_i(s, a | \theta^{Q_i}) \right], \quad (8)$$

where $\mathbb{E}(\cdot)$ denotes the expectation operator, θ^{Q_i} is the parameter of Q_i . As indicated in Fig. 3, the parameter is updated (Update 1) according to:

$$\theta^{Q_i} \leftarrow \theta^{Q_i} - \alpha \nabla_{\theta^{Q_i}} L(\theta^{Q_i}), \quad (9)$$

where α is the learning rate of the critic network. Moreover, TD3 reduces instability by postponing updates to its actor network relative to its critic. In practice, the critic is updated frequently to accurately capture the dynamics of environment, while the actor is adjusted only after a fixed number of critic updates. This delay minimizes error propagation and keeps the policy changes smoother, resulting in more stable and reliable learning outcomes. The actor network is updated (Update 2) with parameters using gradient ascent of the objective function as:

$$J_Q(\theta^\mu) = \mathbb{E} [Q_1(s, \mu(s))], \quad (10)$$

$$\nabla_{\theta^\mu} J_Q(\theta^\mu) = \mathbb{E} [\nabla_a Q_1(s, a | \theta^{Q_1}) \nabla_{\theta^\mu} \mu(s | \theta^\mu)], \quad (11)$$

$$\theta^\mu \leftarrow \theta^\mu + \beta \nabla_{\theta^\mu} J_Q(\theta^\mu), \quad (12)$$

where θ^μ is the parameter of actor network, $J(\theta^\mu)$ is the defined objective function, $\nabla_{\theta^\mu} J(\theta^\mu)$ is the gradient of objective function with respect to θ^μ , β is the learning rate of the policy network. As denoted as Soft update 3 in the Fig. 3, the parameters of the target critic and the target actor network are updated as:

$$\theta_i^{Q'} \leftarrow \tau \theta_i^Q + (1 - \tau) \theta_i^{Q'}, \quad (13)$$

$$\theta^{\mu'} \leftarrow \tau \theta^\mu + (1 - \tau) \theta^{\mu'}, \quad (14)$$

where the soft updating factor is denoted by τ and $\theta^{\mu'}$ is the parameter of the target actor network.

Furthermore, TD3 introduces truncated normal noise to the actions produced by the target policy network, mitigating the trade-off between bias and variance. This added noise serves as a form of regularization, which helps prevent overestimation

of Q-values and reduces the risk of overfitting. By smoothing the target updates in this manner, TD3 achieves more stable learning and enhances the reliability of policy evaluation:

$$a_{t+1} \leftarrow \mu'(s' | \theta^{\mu'}) + \varepsilon, \quad \varepsilon \sim \text{clip}(\mathcal{N}(0, \sigma^2), a_{\min}, a_{\max}), \quad (15)$$

where $a_{\min}$ and $a_{\max}$ define the valid range of actions, representing the lower and upper bounds of the battery charging current, respectively.

1) *Battery application:* For the TD3-based fast charging algorithm, the battery SoC and voltage are chosen as the state variables:

$$s(t) = [V(t), \text{SoC}(t)]^\top. \quad (16)$$

In deployment, SoC can be provided by an onboard estimator, and recent work based on adaptive cubature Kalman filters has demonstrated strong robustness to non-Gaussian disturbances, making such estimators suitable for production BMSs [30]. Similarly, SoH, which is crucial for determining the voltage constraint in our framework, can be accurately estimated using advanced data-driven techniques [31], [32]. Currently, the multi-cell monitor integrated circuits used in EVs can provide millivolt-level cell-voltage accuracy. The production BMSs already measure cell voltage, pack current, and temperature as primary inputs. At the same time, our proposed TD3 controller is robust to moderate sensor noise due to features like target-policy smoothing and twin-critic minimization. These points make our proposed method applicable in real-time deployment without requiring additional specialized sensors.

The action of the agent is the charge current value:

$$a(t) = I(t), \quad s.t. \quad (4). \quad (17)$$

With the established voltage-to-SoH mapping and the TD3 algorithm described above, we define the reward function to address the optimization problem in (5), achieving fast charging while minimizing battery degradation:

$$r(t) = r_{\text{SoC}}(t) + r_{\text{vol}}(t) + r_{\text{smooth}}(t). \quad (18)$$

To encourage fast charging, the SoC-related reward is defined as:

$$r_{\text{SoC}}(t) = \lambda_{\text{SoC}} |\text{SoC}^* - \text{SoC}(t)|, \quad (19)$$

where λ_{SoC} is a weighting factor that quantifies the importance of rapidly reaching the target SoC. To mitigate risks associated with lithium plating, the terminal voltage is constrained by the SoH-dependent upper bound, $V_{\text{max}}(\text{SoH})$:

$$r_{\text{vol}}(t) = \begin{cases} \lambda_{\text{vol}} |V(t) - V_{\text{max}}(\text{SoH})|, & V(t) > V_{\text{max}}(\text{SoH}), \\ 0, & \text{otherwise,} \end{cases} \quad (20)$$

where λ_{vol} is a weighting factor representing the importance of adhering to the voltage limit.

Additionally, to encourage gradual variations in the charging current and prevent abrupt control actions that may stress the battery, we introduce a smoothness term:

$$r_{\text{smooth}}(t) = \lambda_{\text{smooth}} |a(t) - a(t-1)|, \quad (21)$$

where λ_{smooth} is a tuning parameter that penalizes rapid changes in the control input.

It is important to note that, despite the value of SoH being used to update the voltage constraints, it does not appear as a direct term in the reward function. Instead, the influence of SoH is embedded within the dynamic adjustment of V_{max} and the value of V_{max} ensuring that the reinforcement learning policy naturally adapts to the aging mechanism of battery without penalizing the agent explicitly for changes in SoH.

Having established the design of the TD3-based controller, we now proceed to evaluate its performance in the next section.

IV. HEALTH-AWARE FAST CHARGING RESULTS

In this section, we aim to evaluate the proposed RL-based charging strategy using high-fidelity simulations and compare it against widely adopted charging control strategies, such as CC-CV and CC-COP control. To evaluate the effectiveness of incorporating the $V_{\text{cutoff}}\text{--SoH}$ mapping described in Section III, we enhanced the CC-CV method in which the voltage value during the constant-voltage stage is dynamically adjusted based on the SoH-dependent mapping. This modified approach is termed as the CC-CV with varying voltage (CC-CV-V) method, which aims to account for the SoH during battery charging. Furthermore, we introduce another RL benchmark, termed V-constrained RL. This controller utilizes the same TD3 architecture and training approach as our proposed method but operates with a fixed voltage constraint at $V_{\text{max}} = 4.15$ V throughout the lifespan of the battery, representing a more conventional RL charging approach. By comparing our method against V-constrained RL, we can directly quantify the benefits of dynamically adapting the voltage related reward based on SoH.

Additionally, we investigate the constant current–constant overpotential (CC-COP) method as a benchmark controller, using two reference values for the side-reaction overpotential to reflect different charging scenarios. For a conservative strategy ($\eta_{\text{side}}^* = 0.01$ V), a small positive safety margin is applied to stringently avoid lithium plating by accounting for possible dynamic undershoots in the control loop. Motivated by experimental evidence that graphite anodes can tolerate negative voltages before lithium plating occurs [33], we also test a small negative reference ($\eta_{\text{side}}^* = -0.05$ V) to examine

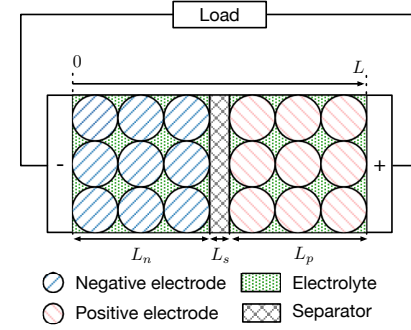


Fig. 4: The internal structure of a lithium-ion battery cell (modified from [14]).

the trade-off between accelerated charging and battery degradation. These variants are denoted CC-COP-slow and CC-COP-fast, respectively, and are used for comparative analysis.

The proposed approach does not require a physical model for implementation as long as the state and reward can be obtained. However, in the context of health-aware battery charging, conducting full-lifetime experiments to collect sufficient data is highly time-consuming. As the primary aim of this work is to introduce and validate an RL-based charging framework, we employ a high-fidelity single-particle model with electrolyte dynamics for training and performance demonstration. The details of the model are provided below.

A. Battery environment for health-aware fast charging

1) *Single particle model with electrolyte*: To provide a reliable and computationally efficient simulation environment for training and evaluating the proposed RL-based charging strategy, this work adopts the SPMe. While the DFN model, or called pseudo-two-dimensional (P2D) model, is a widely used and high-fidelity framework for simulating lithium-ion battery behavior [34], it involves a large number of parameters and intensive computational costs. Moreover, many of its parameters are difficult to measure and calibrate accurately, limiting its practicality in control-oriented applications [35]. In contrast, the SPMe offers a simplified yet sufficiently accurate representation by reducing both the parameter space and computational demand [13]. These features make the SPMe particularly suitable for simulating long-term battery performance and degradation, which are essential considerations in this study.

The structure of a lithium-ion battery is illustrated in Fig. 4, where the negative electrode, separator, and positive electrode have thicknesses L_n , L_s , and L_p , respectively. The SPMe model approximates each electrode as a single spherical particle, assuming that the solid-phase concentration is uniformly distributed along the thickness direction, denoted by x . Under this assumption, the diffusion process inside the spherical particle is governed by

$$\frac{\partial c_{s,i}}{\partial t}(r, t) = \frac{D_{s,i}}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial c_{s,i}}{\partial r}(r, t) \right), \quad i \in \{n, p\}, \quad (22)$$

where r represents the distance from the center of the spherical particle, t is time, $c_{s,i}$ is the solid-phase lithium-ion concentra-

tion, and $D_{s,i}$ is its diffusivity. The electrolyte concentration in all domains are defined as $c_{e,n}$, $c_{e,sep}$, and $c_{e,p}$.

The electric potential in each electrode i is given by

$$\phi_{s,i}(x,t) = \eta_i(t) + \phi_{e,i}(x,t) + U_i(c_{ss,i}(t)) + FR_{f,i}j_{n,i}(t), \quad (23)$$

where η_i is the reaction overpotential, $\phi_{e,i}$ is the potential in the electrolyte phase, and U_i is the open-circuit potential, which is a function of the solid-phase lithium-ion concentration at the particle surface, $c_{ss,i}$. The final term, $FR_{f,i}j_{n,i}(t)$, represents the ohmic potential drop across the SEI film, where F is Faraday's constant, $R_{f,i}$ is the resistance of the SEI film, and $j_{n,i}$ is the molar pore-wall flux of lithium ions.

The nonlinear output of terminal voltage $V(t)$ is computed as:

$$V(t) = \phi_{s,p}(L,t) - \phi_{s,n}(0,t). \quad (24)$$

2) *Battery degradation mechanisms*: Battery performance and lifetime are limited by several degradation mechanisms [36]. In this study, two dominant degradation processes, namely the SEI layer growth and lithium plating are considered. The idea is that suppressing them during fast charging may significantly extend battery lifetime.

The SEI layer forms on the anode particle surface and significantly affects the lifetime of lithium-based batteries. SEI growth is primarily a diffusion-limited process, in which solvent molecules react with lithiated graphite. To accurately capture this process, we utilize a two-layer diffusion-limited SEI growth model described in [37]. In this model, the inner and outer SEI layers are assumed to grow simultaneously at the same rate.

Recent studies have indicated that degradation mechanisms, particularly lithium plating, are strongly influenced by the side-reaction overpotential and can significantly limit battery performance and lifespan [38], [39]. This side-reaction overpotential is described as:

$$\eta_{side}(x,t) = \phi_{s,n} - \phi_{e,n} - U_{side}, \quad (25)$$

where U_{side} is the equilibrium potential of the side reaction, assumed to be zero for lithium plating [40].

The initial partial differential-algebraic equation (PDAE) model that couples the SPMe and the aging model can be reformulated and reduced as a DAE model by suitable numerical methods, such as finite difference and spectral methods [41], [42]:

$$\dot{x} = f(x, z, u), \quad (26)$$

$$y = h(x, z, u), \quad (27)$$

$$0 = g(x, z, u), \quad (28)$$

where $x = [c_{s,n}, c_{s,p}, c_e]^T \in \mathbb{R}^{n_x}$ is the state vector, $z = [\phi_{s,n}, \phi_{s,p}, \phi_e, i_{e,n}, i_{e,p}]^T \in \mathbb{R}^{n_z}$ is the algebraic variable vector, $y = [V, \text{SoC}, \text{SoH}]^T$ is the output and $u \triangleq I$ is the applied input current.

B. Experiment and simulation configurations

The open-source PyBaMM software (version 24.5), is employed as the simulation platform [27], in which the SPMe-aging model from Section IV-A is implemented. As detailed

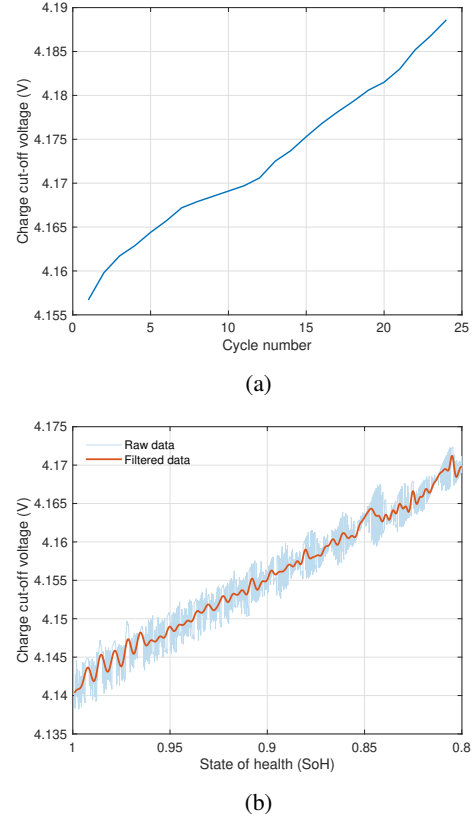


Fig. 5: Extraction and comparison of V_{cutoff} -SoH mapping from (a) experimental data and (b) simulation.

in Section IV-A, the battery model features a single particle model with an electrolyte component, which facilitates efficient and comprehensive analysis of cycle aging. The investigated battery is an LG M50 cell, with electrochemical parameters obtained from [37]. The main degradation mechanisms are characterized by an SEI solvent diffusivity of 2.5×10^{-22} m²/s, a lithium plating kinetic rate constant of 1×10^{-11} m/s, and an irreversible condition for lithium plating.

To obtain the mapping between cut-off voltage and SoH, we first configure the experimental setup as illustrated in Fig. 2. The process relies on a three-electrode cell configuration, such as the EL-CELL PAT-Cell system. This type of cell incorporates a reference electrode in the middle of the separator, which makes the voltage between each current collector and the reference electrode measurable. As a result, the voltage of each individual electrode can be approximately determined. For our specific tests, the positive electrode consists of nickel manganese cobalt (NMC) oxide of type 811, and the negative electrode uses artificial graphite. All assembly was conducted in an argon-filled glove box, and cycling was performed in a temperature-controlled chamber at a constant 25°C. Then, the mappings derived from both experimental measurements and high-fidelity simulation are compared as shown in Fig. 5. Although different battery cells are used in the experiment and simulation, the trend between the charge cut-off voltage and SoH remains consistent across both domains.

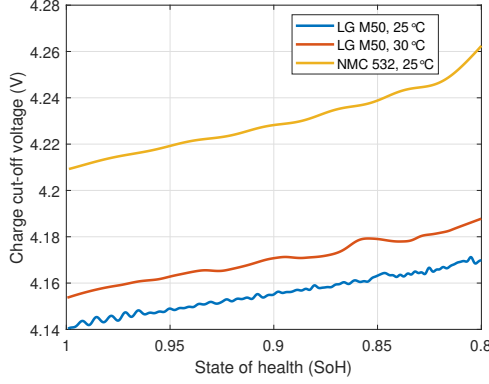


Fig. 6: V_{cutoff} -SoH mapping under different temperature and chemistry conditions based on filtered data.

In practice, the V_{cutoff} -SoH map may be dependent on operating temperature and cell chemistry, as illustrated by the results shown in Fig. 6, which were the simulation results using PyBaMM. For real-world deployment, this implies that a condition-aware prediction model for V_{cutoff} would be necessary, taking SoH, temperature, and chemistry as inputs. However, in applications like EVs, thermal management systems typically narrow the range of temperature excursions, which would help mitigate these variations. For the performance evaluation in this study, we focused on the LG M50 cell at 25°C to align our work with a well-validated set of degradation parameters from [37]. This choice was crucial for ensuring the fundamental fidelity of our simulation environment.

We implement a unified framework to assess various charging strategies across the entire battery lifetime. The details of this process are shown as Protocol 1. This protocol includes an initial stage to obtain the actual battery capacity, followed by repeated testing cycles using different charging controls. The primary objective of this comprehensive battery life testing is to evaluate the effectiveness of these controls in optimizing the charging process and extending the operational lifespan of the battery.

1) *Controller configuration and training results:* The proposed controller employs an actor-critic neural network architecture. The actor and critic networks each contain two fully connected hidden layers with 200 units per layer. The training hyperparameters are summarized in Table I. The penalty coefficients in the proposed controller are chosen as $\lambda_{\text{SoC}} = -2$, $\lambda_{\text{vol}} = -10$, and $\lambda_{\text{smooth}} = -0.5$, respectively. The maximum control input is limited as $I_{\text{max}} = 10\text{A}$. The lifelong TD3 agent was trained on a consumer-grade desktop computer equipped with an Intel i5-13400F CPU, 32 GB of RAM, and an NVIDIA RTX 3060 GPU, running on Windows 11. The entire training process, covering approximately 2000 episodes as presented in our results, required about 22 hours. This demonstrates that the proposed framework can be effectively trained on widely available hardware without needing access to specialized high-performance computing clusters, confirming its feasibility for practical implementation.

The training performance of the proposed charging algorithm is presented in Fig. 7, illustrating the evolution of four

Protocol 1 Battery cycling protocol for controller investigation

- 1: **Initialize:** Discharge the battery to 0% SoC.
- 2: **Charge** to 100% SoC using CC-CV.
- 3: **Discharge** the battery completely to determine actual capacity.
- 4: **repeat**
- 5: **Charge** the battery from 0% SoC to 20% SoC using low-current CC strategy.
- 6: **Rest** for one hour.
- 7: **Apply** the target charging strategy to charge the battery to 80% SoC.
- 8: **Discharge** the battery to 0% SoC.
- 9: **Rest** for one hour.
- 10: **Check** SoH; if SoH falls to SoH_{end} (e.g., 80%), end test.
- 11: **until** battery $\text{SoH} \leq \text{SoH}_{\text{end}}$
- 12: **Evaluate** performance based on total cycles in terms of the average charging time and the equivalent full cycles.

TABLE I: Hyperparameters of the proposed TD3 algorithm

Hyperparameters	Values
Actor network learning rate	0.0001
Critic network learning rate	0.0001
Discounting factor	0.99
Experience replay buffer size	1.0×10^6
Minibatch size	256
Soft update factor	0.005
Delay frequency	1

critical metrics across approximately 2000 episodes. The evaluations were conducted at 20 episode intervals. Fig. 7a shows the cumulative reward, which initially fluctuates significantly around -2000 due to the exploration of highly conservative charging strategies but stabilizes near -60 after approximately 1000 episodes, indicating convergence to a policy effectively reducing charging time while extending battery life. The large initial negative values result from assigning a substantial negative reward (-1000) whenever the charging time exceeds 480 minutes. Fig. 7b depicts the charge cut-off voltage, which initially oscillates around 3.5 V but stabilizes around 4.2 V after 1000 episodes, demonstrating the ability of algorithm to maintain an optimal voltage threshold guided by the V_{cutoff} -SoH mapping. Additionally, Fig. 7c illustrates the convergence behavior of the minimum side-reaction overpotential, showing that its value stabilizes around -0.145V . Fig. 7d demonstrates that, in general, higher cumulative rewards during training correlate with reduced charging times. The charging time ultimately stabilizes near 27 minutes in later episodes.

C. Fast charging results

With the TD3 agent successfully trained and its convergence confirmed as shown in Fig. 7, we now proceed to evaluate its lifelong charging performance. Following the battery cycling protocol as described in Protocol 1, we conducted the battery charging tests using six different strategies: the CC-CV strategy, the CC-CV-V strategy, two CC-COP strategies, the V-constrained RL, and the proposed approach. A finer

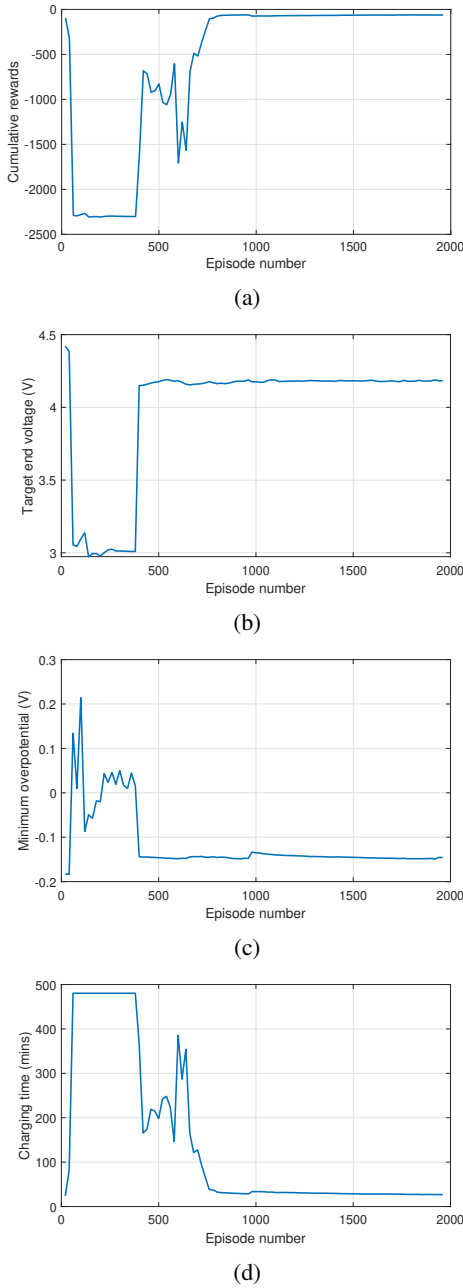


Fig. 7: The training performance of the proposed method: (a) cumulative rewards; (b) maximum end voltage; (c) minimum side-reaction overpotential η_{side} ; (d) charging time.

sampling resolution can better capture the transient dynamics of batteries, but at the expense of computational efficiency. In this study, the primary metrics of interest are battery lifetime and charging time. Considering these factors, the sampling time for the lifelong implementation of all these strategies was set as 20 s. The EFCs and the average charging time for these approaches are summarized in Table II.

1) *Overall analysis on charging time and lifespan:* Among the methods, the CC-COP controllers rely fully on the measurement of side-reaction overpotential, and demonstrate performance variation depending on the tuning of the desired

TABLE II: Battery charging performance based on different methods

	Maximum EFCs	Average charging time (mins)	Relative lifespan extension
CC-CV	572	24.15	-
CC-CV-V	611	24.18	6.82%
CC-COP slow*	1316	36.02	130%
CC-COP fast*	1009	22.40	76.4%
V-constrained RL	664	24.34	16.1%
Proposed method	703	24.12	22.9%

* Requires side-reaction overpotential measurements.

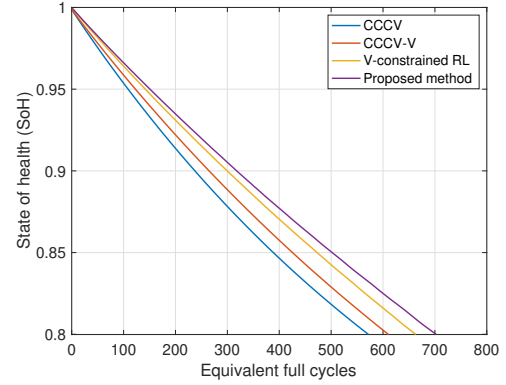


Fig. 8: Comparison of SoH degradation among different charging methods.

side-reaction overpotential reference. The CC-COP-fast strategy achieves the shortest average charging time of 22.40 minutes and significantly extends the battery lifetime compared to CC-CV methods, with 1009 EFCs. In comparison, the CC-COP-slow strategy attains the longest battery lifespan, which is 1316 EFCs, yet requires the longest charging time of 36.02 minutes, reflecting a very conservative strategy.

Different from CC-COP strategies, the proposed approach along with CC-CV, CC-CV-V and V-constrained RL exhibit more practical applicability and deployability. The variation in battery SOH with respect to EFCs for these four methods is presented in Fig. 8. The baseline CC-CV method, despite achieving a competitive charging time of 24.15 minutes, suffers from the shortest battery lifespan with 572 EFCs. This limitation stems from its rigid voltage and current thresholds during the charging process, which fail to adapt to battery degradation dynamics. The CC-CV-V strategy addresses this weakness by incorporating V_{cutoff} -SOH mapping to dynamically adjust the charge cut-off voltage. This modification yields a modest lifespan improvement with 611 EFCs while maintaining near-identical charging speed with 24.14 minutes. However, the 6.8% improvement of EFCs over CC-CV suggests that static parameter mapping alone cannot fully capture complex aging mechanisms. The V-constrained RL benchmark shows a greater lifespan improvement than the CC-CV variants. It achieves 664 EFCs, a 16.1% extension over the baseline, while maintaining a competitive average charging time of 24.34 minutes.

In contrast, the proposed method achieves a significant im-

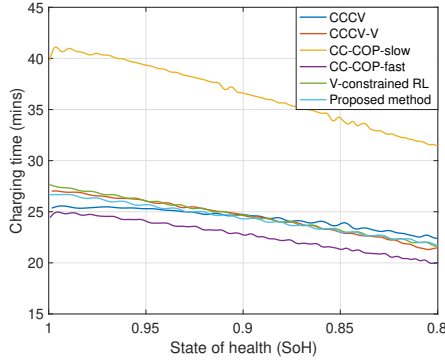


Fig. 9: Charging time based on different charging approaches.

provement in performance, where battery lifespan is extended to 703 EFCs with a marginally reducing charging time to 24.12 minutes. This represents a 22.9% improvement over CC-CV baseline and a notable improvement over the V-constrained RL benchmark. This breakthrough results from its health-aware control architecture, which integrates indirect degradation indicators to dynamically optimize charging protocols.

In summary, compared to CC-COP strategies, the proposed method achieves enhanced practicality for real-world deployment by eliminating the need for specialized overpotential sensors, despite its slightly longer charging time and fewer equivalent full cycles than CC-COP-fast and CC-COP-slow approaches. In contrast to CC-CV and its variants, the proposed approach maintains comparable charging efficiency while largely extending battery lifespan, demonstrating that lifespan enhancement can be achieved without compromising charging speed.

2) *Charging time dynamics over battery aging:* While the aggregate results in Table II provide a clear summary of lifespan and average charging time, it is also crucial to understand how charging dynamics evolve as the battery ages. Fig. 9 compares six charging strategies in terms of charging time versus SoH. Overall, the CC-COP-slow method exhibits the longest charging durations across the SoH range, starting at approximately 40 minutes when the battery is fully healthy and dropping to around 30 minutes when SoH reaches 80%. The charging times of all methods decrease as SoH declines, primarily because less energy or charge capacity is absorbed by the battery for the same SoC change. The CC-CV, CC-CV-V, V-constrained RL, and proposed methods exhibit generally similar trajectories, though CC-CV-V notably shows a substantial peak at the beginning and subsequently decreases more rapidly than the CC-CV and the proposed method. By contrast, the proposed method maintains a relatively stable decrease in charging time compared to the other methods.

3) *Evolution of current and overpotential profiles:* To further investigate why the proposed learning-based method effectively extends the battery lifespan compared to CC-CV and its variants, the battery current and side-reaction overpotential profiles during a single charging process are plotted for two distinct aging stages. Specifically, results from the 1st cycle (new battery) and the 200th cycle (aged battery) are presented in Fig. 10.

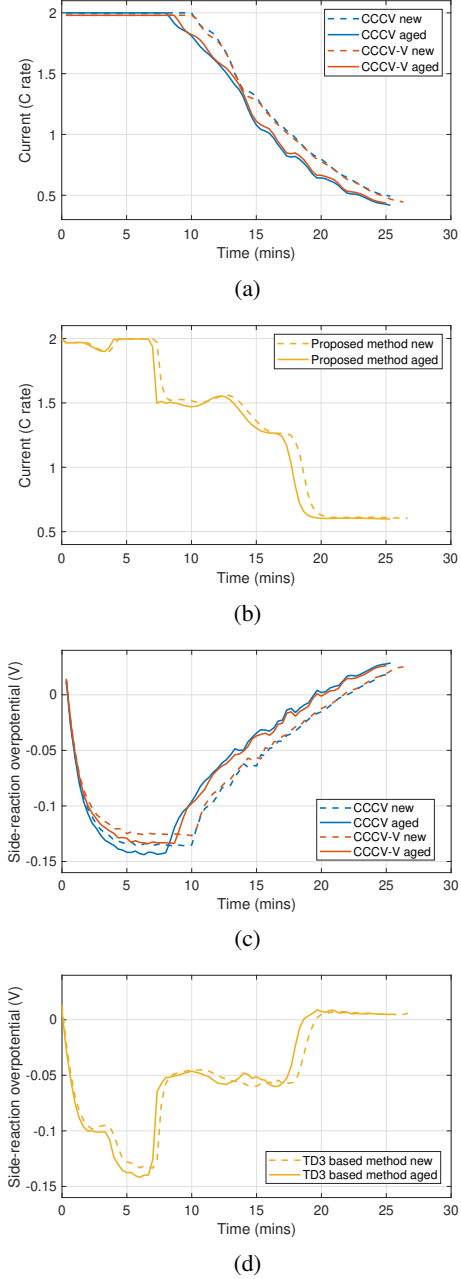


Fig. 10: Current and side-reaction overpotential profiles of the studied cell at different aging levels under different control methods.

Fig. 10 illustrates distinct shifts in side-reaction overpotential profiles across different charging methods when comparing new and aged cells. For aged cells, a clear downward and leftward shift in the side-reaction overpotential curves is observed, indicating that the same charging current results in lower overpotential as the cell ages. This phenomenon is particularly pronounced in the CC-CV and CC-CV-V charging methods. In contrast, the current profiles of the proposed method behave like a multi-stage constant-current scheme that results in smoother current and side-reaction overpotential transitions in both new and aged cells. The TD3-based method significantly

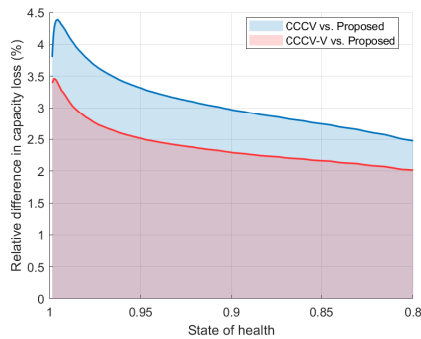


Fig. 11: Relative difference in capacity loss due to lithium plating by comparing CC-CV and CC-CV-V to the proposed method. The positive value in y -axis means the least plated lithium triggered by the proposed method.

reduces the difference in side-reaction overpotential between new and aged cells, indicating its ability to attenuate the effects of battery aging.

Regarding overpotential levels, the proposed approach only remains at a relatively low negative value for a short period before moving the side-reaction overpotential closer to zero. By contrast, the side-reaction overpotential under CC-CV and its variants methods remains below -0.05 V for most of the charge and lasts for nearly 15 minutes. This prolonged negative overpotential may thermodynamically favor lithium plating, thereby increasing the risk of degradation and reducing long-term battery health compared to the proposed method.

These results not only facilitate a comparison between the different controllers but also highlight the discrepancies in performance of the same controller at different stages of battery aging. Such insights are especially meaningful for researchers and engineers seeking to optimize fast charging strategies.

4) *Capacity loss due to lithium plating*: Finally, we analyze the relative difference in capacity loss for CC-CV and CC-CV-V, compared to the proposed method, as a function of SoH, as shown in Fig. 11. This difference is expressed as a percentage of the baseline capacity loss for each strategy. Positive values indicate that the proposed method yields a lower capacity loss than conventional strategies, highlighting its effectiveness in extending battery life.

V. CONCLUSION

In this paper, we present a novel health-aware fast charging strategy for lithium-ion batteries, based on a deep learning approach. Unlike conventional charging methods that primarily focus on minimizing charging time, our proposed approach explicitly considers the trade-off between fast charging and extending battery life. By utilizing a mapping between the charge end voltage and the SoH of the battery from a constant current and constant overpotential control, the proposed method incorporates SoH-dependent voltage into the optimal control decision-making process, effectively mitigating adverse degradation phenomena.

To demonstrate the effectiveness of the proposed method, we utilized a high-fidelity single particle model with electrolyte implemented in PyBaMM. This model served as the test environment, capturing realistic degradation behaviors and allowing for robust evaluation across a full life-cycle simulation. Comparisons with benchmark controllers demonstrate that the TD3-based policy successfully reduces overall degradation without significantly compromising charge speed, showing clear promise for practical implementation. Moreover, we believe that machine learning methods such as transfer learning can significantly reduce the experimental workload required to generate the V_{cutoff} -SoH maps for new battery types, which can further enhance the applicability of this approach.

ACKNOWLEDGMENTS

The authors would like to thank Dr Yang Li from the Chalmers University of Technology for his assistance with potential model simulation and setup in MATLAB.

REFERENCES

- [1] L. Zhang, C. Sun, G. Cai, and L. H. Koh, "Charging and discharging optimization strategy for electric vehicles considering elasticity demand response," *ETransportation*, vol. 18, p. 100262, 2023.
- [2] J. B. Goodenough and K.-S. Park, "The Li-ion rechargeable battery: A perspective," *J. Am. Chem. Soc.*, vol. 135, no. 4, pp. 1167–1176, 2013.
- [3] A. Ahmad, J. Meyboom, P. Bauer, and Z. Qin, "Techno-economic analysis of energy storage systems integrated with ultra-fast charging stations: A dutch case study," *eTransportation*, p. 100411, 2025.
- [4] G. E. Blomgren, "The development and future of lithium ion batteries," *J. Electrochem. Soc.*, vol. 164, no. 1, p. A5019, 2016.
- [5] H. Rahimi-Eichi, U. Ojha, F. Baronti, and M.-Y. Chow, "Battery management system: An overview of its application in the smart grid and electric vehicles," *IEEE Ind. Electron. Mag.*, vol. 7, no. 2, pp. 4–16, 2013.
- [6] M. Broussely, P. Biensan, F. Bonhomme, P. Blanchard, S. Herreyre, K. Nechev, and R. Staniewicz, "Main aging mechanisms in Li ion batteries," *J. Power Sources*, vol. 146, no. 1–2, pp. 90–96, 2005.
- [7] P. M. Attia, A. Grover, N. Jin, K. A. Severson, T. M. Markov, Y.-H. Liao, M. H. Chen, B. Cheong, N. Perkins, Z. Yang, *et al.*, "Closed-loop optimization of fast-charging protocols for batteries with machine learning," *Nature*, vol. 578, no. 7795, pp. 397–402, 2020.
- [8] Y. Li, T. Wik, Y. Huang, and C. Zou, "Nonlinear model inversion-based output tracking control for battery fast charging," *IEEE Trans. Control Syst. Technol.*, vol. 32, no. 1, pp. 225–240, 2023.
- [9] M. Yuan, A. Burman, and C. Zou, "Robust model predictive control of fast lithium-ion battery pretreatment for safe recycling," *arXiv preprint arXiv:2503.11857*, 2025.
- [10] Z. Khalik, M. Donkers, J. Sturm, and H. J. Bergveld, "Parameter estimation of the doyle–fuller–newman model for lithium-ion batteries by parameter normalization, grouping, and sensitivity analysis," *J. Power Sources*, vol. 499, p. 229901, 2021.
- [11] P. W. Northrop, B. Suthar, V. Ramadesigan, S. Santhanagopalan, R. D. Braatz, and V. R. Subramanian, "Efficient simulation and reformulation of lithium-ion battery models for enabling electric transportation," *J. Electrochem. Soc.*, vol. 161, no. 8, p. E3149, 2014.
- [12] G. Hwang, N. Sitapure, J. Moon, H. Lee, S. Hwang, and J. S.-I. Kwon, "Model predictive control of lithium-ion batteries: Development of optimal charging profile for reduced intracycle capacity fade using an enhanced single particle model (SPM) with first-principled chemical/mechanical degradation mechanisms," *Chem. Eng. J.*, vol. 435, p. 134768, 2022.
- [13] S. J. Moura, F. B. Argomedeo, R. Klein, A. Mirtabatabaei, and M. Krstic, "Battery state estimation for a single particle model with electrolyte dynamics," *IEEE Trans. Control Syst. Technol.*, vol. 25, no. 2, pp. 453–468, 2016.
- [14] S. G. Marquis, V. Sulzer, R. Timms, C. P. Please, and S. J. Chapman, "An asymptotic derivation of a single particle model with electrolyte," *J. Electrochem. Soc.*, vol. 166, no. 15, p. A3693, 2019.

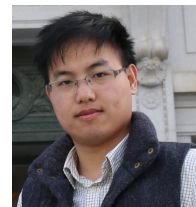
- [15] W. Huang, Y. Ye, H. Chen, R. A. Vilá, A. Xiang, H. Wang, F. Liu, Z. Yu, J. Xu, Z. Zhang, *et al.*, "Onboard early detection and mitigation of lithium plating in fast-charging batteries," *Nat. Commun.*, vol. 13, no. 1, p. 7091, 2022.
- [16] N. Wassiliadis, J. Kriegler, K. A. Gamra, and M. Lienkamp, "Model-based health-aware fast charging to mitigate the risk of lithium plating and prolong the cycle life of lithium-ion batteries in electric vehicles," *J. Power Sources*, vol. 561, p. 232586, 2023.
- [17] X. Yang, Z. Wei, and L. Du, "Constant overpotential fast charging for lithium-ion battery with twin delayed DDPG algorithm," in *2022 IEEE Transportation Electrification Conference & Expo (ITEC)*, 2022, pp. 623–628.
- [18] Y. Lu, X. Han, Y. Li, X. Li, and M. Ouyang, "Health-aware fast charging for lithium-ion batteries: Model predictive control, lithium plating detection, and lifelong parameter updates," *IEEE Trans. Ind. Appl.*, 2024.
- [19] H. E. Perez, X. Hu, S. Dey, and S. J. Moura, "Optimal charging of Li-ion batteries with coupled electro-thermal-aging dynamics," *IEEE Trans. Veh. Technol.*, vol. 66, no. 9, pp. 7761–7770, 2017.
- [20] Y. Zhang, T. Wik, J. Bergström, and C. Zou, "Machine learning-based lifelong estimation of lithium plating potential: A path to health-aware fastest battery charging," *Energy Storage Mater.*, vol. 74, p. 103877, 2025.
- [21] S. Zhao, X. Sun, Y. An, Z. Guo, C. Li, Y. Xu, Y. Li, Z. Li, X. Zhang, K. Wang, *et al.*, "Lithium plating accurate detection of Lithium-Ion capacitors upon high-rate charging," *Green Energy Intell. Transp.*, p. 100268, 2025.
- [22] S. Park, A. Pozzi, M. Whitmeyer, H. Perez, A. Kandel, G. Kim, Y. Choi, W. T. Joe, D. M. Raimondo, and S. Moura, "A deep reinforcement learning framework for fast charging of Li-ion batteries," *IEEE Trans. Transp. Electr.*, vol. 8, no. 2, pp. 2770–2784, 2022.
- [23] Z. Wei, Z. Quan, J. Wu, Y. Li, J. Pou, and H. Zhong, "Deep deterministic policy gradient-DRL enabled multiphysics-constrained fast charging of lithium-ion battery," *IEEE Trans. Ind. Electron.*, vol. 69, no. 3, pp. 2588–2598, 2021.
- [24] Y. Gao, J. Zhu, D. Shi, and X. Zhang, "A model-based battery dataset recovery method considering cell aging in real-world electric vehicles," *IEEE Trans. Ind. Inform.*, vol. 20, no. 5, pp. 7904–7914, 2024.
- [25] Y. Gao, X. Zhang, C. Zhu, and B. Guo, "Global parameter sensitivity analysis of electrochemical model for lithium-ion batteries considering aging," *IEEE/ASME Trans. Mechatron.*, vol. 26, no. 3, pp. 1283–1294, 2021.
- [26] Y. Hao, Q. Lu, X. Wang, and B. Jiang, "Adaptive model-based reinforcement learning for fast charging optimization of lithium-ion batteries," *IEEE Trans. Ind. Inform.*, 2023.
- [27] V. Sulzer, S. G. Marquis, R. Timms, M. Robinson, and S. J. Chapman, "Python battery mathematical modelling (PyBaMM)," *J. Open Res. Softw.*, vol. 9, no. 1, 2021.
- [28] S. Fujimoto, H. Hoof, and D. Meger, "Addressing function approximation error in actor-critic methods," in *Int. Conf. Mach. Learn. (ICML)*. PMLR, 2018, pp. 1587–1596.
- [29] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," in *Proceedings of the 4th International Conference on Learning Representations (ICLR)*, 2016.
- [30] C. Chen, Q. Zhang, W. Liao, F. Zhu, M. Li, and H. Wu, "State of charge estimation for lithium-ion batteries using an adaptive cubature kalman filter based on improved generalized minimum error entropy criterion," *Green Energy Intell. Transp.*, p. 100292, 2025.
- [31] C. Zhang, L. Tu, Z. Yang, B. Du, Z. Zhou, J. Wu, and L. Chen, "A CMMOG-based lithium-battery SOH estimation method using multi-task learning framework," *J. of Energy Storage*, vol. 107, p. 114884, 2025.
- [32] C. Zhang, L. Tu, Z. Zhou, S. Chen, J. Wu, and L. Chen, "Estimating lithium-ion battery health using hybrid attention networks and multi-source data," *IEEE Sens. J.*, 2024.
- [33] T. Gao, Y. Han, D. Fraggedakis, S. Das, T. Zhou, C.-N. Yeh, S. Xu, W. C. Chueh, J. Li, and M. Z. Bazant, "Interplay of lithium intercalation and plating on a single graphite particle," *Joule*, vol. 5, no. 2, pp. 393–414, 2021.
- [34] L. Xia, E. Najafi, Z. Li, H. Bergveld, and M. Donkers, "A computationally efficient implementation of a full and reduced-order electrochemistry-based model for Li-ion batteries," *Appl. Energy*, vol. 208, pp. 1285–1296, 2017.
- [35] Z. Khalik, M. Donkers, and H. J. Bergveld, "Model simplifications and their impact on computational complexity for an electrochemistry-based battery modeling toolbox," *J. Power Sources*, vol. 488, p. 229427, 2021.
- [36] H. Wang, Y. Song, X. Sun, S. Mo, C. Chen, and J. Wang, "Onboard in-situ warning and detection of Li plating for fast-charging batteries with deep learning," *Energy Storage Mater.*, vol. 71, p. 103585, 2024.
- [37] S. E. O'Kane, W. Ai, G. Madabattula, D. Alonso-Alvarez, R. Timms, V. Sulzer, J. S. Edge, B. Wu, G. J. Offer, and M. Marinescu, "Lithium-ion battery degradation: how to model it," *Phys. Chem. Chem. Phys.*, vol. 24, no. 13, pp. 7909–7922, 2022.
- [38] N. A. Chaturvedi, R. Klein, J. Christensen, J. Ahmed, and A. Kojic, "Algorithms for advanced battery-management systems," *IEEE Control Syst. Mag.*, vol. 30, no. 3, pp. 49–68, 2010.
- [39] A. Tomaszewska, Z. Chu, X. Feng, S. O'kane, X. Liu, J. Chen, C. Ji, E. Endler, R. Li, L. Liu, *et al.*, "Lithium-ion battery fast charging: A review," *eTransportation*, vol. 1, p. 100011, 2019.
- [40] P. Arora, M. Doyle, and R. E. White, "Mathematical modeling of the lithium deposition overcharge reaction in lithium-ion batteries using carbon-based negative electrodes," *J. Electrochem. Soc.*, vol. 146, no. 10, p. 3543, 1999.
- [41] Y. Li, D. Karunathilake, D. M. Vilathgamuwa, Y. Mishra, T. W. Farrell, S. S. Choi, and C. Zou, "Model order reduction techniques for physics-based lithium-ion battery management: A survey," *IEEE Ind. Electron. Mag.*, vol. 16, no. 3, pp. 36–51, 2021.
- [42] Y. Huang, C. Zou, Y. Li, and T. Wik, "MINN: Learning the dynamics of differential-algebraic equations and application to battery modeling," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2024.



Meng Yuan received the B.E. degree in automation and the M.Sc. degree in control theory and control engineering from Northeastern University, China, in 2013 and 2015, respectively, and the Ph.D. degree in electrical and electronic engineering from the University of Melbourne, Australia, in 2020. He was a Research Fellow at the Rehabilitation Research Institute of Singapore, Nanyang Technological University, Singapore. He joined the College of Electrical Engineering and Automation, Fuzhou University, China in 2023, as a Faculty Member.

He is currently a Researcher in the Department of Electrical Engineering at Chalmers University of Technology, Sweden. His research interests include modeling, learning-based control and model predictive control for batteries, energy systems, mechatronics, and robotics.

Dr. Yuan is a recipient of the EU Marie Skłodowska-Curie Action Post-doctoral Fellowship, and he was a recipient of the Keynote Paper Award at the 25th Chinese Process Control Conference.



Changfu Zou (Senior Member, IEEE) received the Ph.D. degree in automation and control engineering from the University of Melbourne, Melbourne, VIC, Australia, in 2017.

He was a Visiting Student Researcher at the University of California, Berkeley, Berkeley, CA, USA, from 2015 to 2016. He has been with the Automatic Control unit, Chalmers University of Technology, Gothenburg, Sweden, since 2017, progressing from a Researcher to an Assistant Professor, then an Associate Professor, and is currently a Professor.

His research focuses on advanced modeling and automatic control of energy storage systems, particularly batteries.

Dr. Zou has been awarded several prestigious grants from the European Commission and Swedish national agencies and has hosted five researchers to achieve the Marie Skłodowska-Curie Fellowships. He serves as an Associate Editor for several journals such as IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY and IEEE TRANSACTIONS ON TRANSPORTATION ELECTRIFICATION.