

Article

Fast-Charging Optimization Method for Lithium-Ion Battery Packs Based on Deep Deterministic Policy Gradient Algorithm

Zhi Zhang ¹, Taijun Guo ¹, Yefeng Liu ², Xinfu Pang ^{3,*} and Zedong Zheng ⁴¹ School of Mechanical Engineering and Automation, Shenyang Institute of Technology, Fushun 113122, China; zhangzhi@situ.edu.cn (Z.Z.); guotaijun@situ.edu.cn (T.G.)² School of Automation and Electrical Engineering, Linyi University, Linyi 276005, China; liuyefeng@lyu.edu.cn³ Key Laboratory of Energy Saving and Controlling in Power System of Liaoning Province, Shenyang Institute of Engineering, Shenyang 110136, China⁴ School of Computing and Mathematical Sciences, University of Leicester, Leicester LE1 7RH, UK; zz288@leicester.ac.uk

* Correspondence: pangxf@sie.edu.cn

Abstract: Fast-charging technology for lithium-ion batteries is of great significance in reducing charging time and enhancing user experience. However, during fast charging, the imbalance among battery cells can affect the overall performance and available capacity of the battery pack. Moreover, the charging efficiency not only is limited by the battery technology itself but is also closely related to the optimization of the charging strategy. To address the optimization of fast charging for lithium-ion batteries, this paper proposes a method based on deep reinforcement learning. First, a deep reinforcement learning charging optimization model is constructed, aiming to minimize charging time and SOC balancing cost, with constraints on battery voltage, temperature, SOC, and SOH. The model employs the deep deterministic policy gradient (DDPG) algorithm integrated with reward centralization and entropy regularization mechanisms, aiming to dynamically adjust the charging current to achieve an optimal balance between fast charging and battery health. Experimental results indicate that the proposed method enhances charging efficiency, contributes to extending battery life, and supports the safety of the charging process. Compared to the traditional constant-current constant-voltage (CCCV) strategy, the improved DDPG strategy reduces the total charging time by 60 s and the balancing time from 540 s to 470 s. Furthermore, compared to the basic DDPG method, the proposed algorithm shows a clear advantage in charging efficiency.

Keywords: charging time; deep deterministic policy gradient (DDPG) algorithm; deep reinforcement learning; lithium-ion battery fast charging; SOC balancing cost



Academic Editor: King Jet Tseng

Received: 5 April 2025

Revised: 11 May 2025

Accepted: 15 May 2025

Published: 17 May 2025

Citation: Zhang, Z.; Guo, T.; Liu, Y.; Pang, X.; Zheng, Z. Fast-Charging Optimization Method for Lithium-Ion Battery Packs Based on Deep Deterministic Policy Gradient Algorithm. *Batteries* **2025**, *11*, 199. <https://doi.org/10.3390/batteries11050199>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Literature Review

With the rapid development of renewable energy and the continuous expansion of the electric vehicle (EV) market, lithium-ion batteries (LIBs) have become a key focus in energy storage optimization. A critical function of the battery management system (BMS) is battery balancing, which ensures the efficient operation and extended lifespan of battery packs by maintaining similar voltage or state of charge (SOC) levels across individual cells [1].

Fast-charging technology not only significantly reduces the charging time for EVs, improving user experience, but also plays a crucial role in grid peak shaving and energy storage systems [2]. However, during fast charging, the increase in battery temperature and

the imbalance in electrochemical reactions can lead to reduced battery life and increased safety risks [3]. Therefore, how to reduce charging time while maintaining battery state consistency has become a critical technical challenge.

Traditional charging strategies typically use battery packs composed of identical cells. However, inherent parameter differences among cells [4] can lead to imbalances during charging, affecting the overall performance and available capacity of the battery pack. Consequently, effective balancing strategies are essential to mitigate such cell inconsistencies and maintain operational homogeneity across the battery system. Existing balancing techniques are mainly divided into passive and active methods [5]. Passive balancing, the most common and economical method in industry, dissipates excess energy from cells with higher SOC or voltage as heat through resistors [6]. Although cost-effective and simple, passive balancing suffers from limitations including extended charging durations and heat generation resulting from energy loss [7].

Active battery balancing primarily encompasses three implementation approaches: rule-driven, model-driven, and intelligent optimization methods [8–10]. Among these, rule-driven approaches employ predefined algorithmic frameworks and heuristic control strategies to dynamically adjust critical charging parameters, including but not limited to current thresholds and voltage limits, based on battery characteristics and actual operating conditions [11,12]. These model-free strategies are widely used in practice, including the constant-current constant-voltage (CCCV) charging protocol and its variants, such as multistage constant current (MCC) [13], multistage CCCV [14], and boost charging [15]. Although existing methods are straightforward to implement and computationally efficient, they fall short of achieving optimal performance in terms of charging speed, safety, and battery lifespan due to a lack of in-depth understanding of battery dynamics and physical constraints. Furthermore, traditional methods often fail to adequately account for variations between individual battery cells and the effects of aging, resulting in a lack of precise control during the charging process, which can lead to uneven temperature rise and reduced efficiency.

Model-based approaches first leverage various battery models to capture battery dynamics and estimate state of charge (SOC) [16–18]. Subsequently, specific objective functions and constraints are formulated by incorporating multiple performance metrics, including charging speed, degradation rate, and thermal characteristics. However, model-based strategies are highly sensitive to model accuracy and require significant computational resources.

Recent years have witnessed significant advancements in intelligent charging algorithms [19]. As a powerful optimization tool, reinforcement learning (RL) has demonstrated considerable potential in charging strategy optimization due to its strong adaptability and suitability for complex dynamic systems [20]. Through continuous interaction with battery systems, RL algorithms dynamically adjust charging parameters, effectively improving charging speeds while mitigating battery degradation risks [21–23]. Despite these achievements, several critical challenges remain unresolved. For instance, Hao et al. [24] proposed a data-driven charging control method based on proximal policy optimization (PPO), which derives the optimal charging strategy through offline training. However, this method does not fully consider safety constraints during the charging process, such as overheating and overvoltage, which can impact its reliability in practical applications. Wang et al. introduced a reinforcement learning approach that strikes an intelligent balance between fast charging and battery lifespan, optimizing the charging configuration and extending the lifespan of lithium-ion batteries [25]. Although this method achieves a balance between fast charging and battery health to some extent, it primarily focuses on optimizing the overall charging process and pays less attention to the consistency of individual battery

cells. In particular, in complex charging environments or unstable battery conditions, these methods may not effectively ensure battery safety and longevity.

Building upon this, the present study introduces an innovative charging strategy for lithium-ion battery packs, leveraging the deep deterministic policy gradient (DDPG) algorithm to enable intelligent control and adaptive optimization of the charging process. The proposed strategy addresses critical challenges such as rapid charging, maintaining cell-to-cell consistency, and dynamically adhering to thermal and electrochemical constraints. The approach aims to optimize charging efficiency and battery performance while safeguarding the overall safety of the charging process.

1.2. Motivation and Contributions

The existing research primarily suffers from the following issues: The inherent parameter variations among battery cells gradually intensify over time, while current balancing mechanisms lack dynamic correction capabilities, leading to a decline in the overall performance and available capacity of the battery pack. Additionally, the charging efficiency of battery packs is constrained by suboptimal charging strategies, as existing methods struggle to simultaneously achieve SOC balancing and meet fast-charging demands, thereby compromising system performance. Inspired by the above literature, this paper proposes a novel deep deterministic policy gradient (DDPG) algorithm model and thoroughly validates the proposed method. The main contributions of this paper are as follows.

- (1) Combining the electrothermal battery model with a novel continuous DRL algorithm to simulate the dynamic behavior of lithium-ion battery packs, enabling more accurate charging control;
- (2) Innovatively constructing a cost function that integrates charging time and battery state (SOC) balancing, with appropriate penalty terms to minimize charging time while ensuring battery state consistency during charging, thereby avoiding overcharging and uneven charging risks;
- (3) Proposing a DRL-based charging method using the DDPG algorithm, integrated with reward centralization and entropy regularization mechanisms. This innovative combination dynamically adjusts the charging current to optimize the balance between fast charging and battery health.

The paper is organized as follows. Section 2 presents the electrothermal battery model. Section 3 formulates the charging optimization problem with its objectives and constraints. Section 4 details the improved DDPG-based charging optimization method. Section 5 analyzes the simulation results and compares performance with conventional approaches. Section 6 summarizes this article.

2. Electrothermal Model of Lithium-Ion Battery

In modeling lithium-ion batteries (LIBs), it is crucial to consider the coupling between electrical characteristics and thermal effects, as temperature changes significantly impact battery performance and lifespan. Therefore, the electrothermal coupling model is key to describing battery dynamic characteristics and optimizing battery management strategies.

This paper adopts a second-order RC circuit model to describe the electrical behavior of the battery [26], combined with a thermal model to simulate heat generation and transfer during charging and discharging, comprehensively reflecting the electrical and thermal characteristics of the battery. Although higher-order and electrochemical models offer improved accuracy, they also introduce a greater number of parameters and increased computational complexity, which significantly raises the computational burden. In contrast, the second-order RC model achieves a favorable balance between modeling fidelity and computational feasibility, as shown in Figure 1.

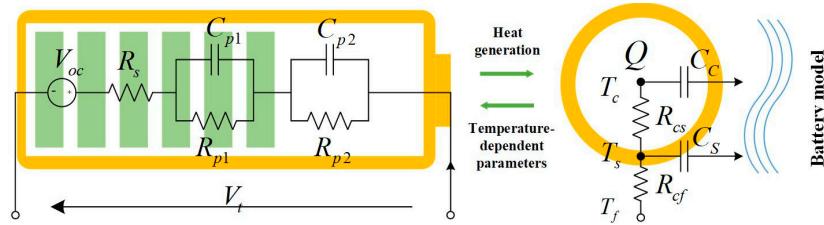


Figure 1. Electro-thermal model of the lithium-ion battery.

2.1. Electrical Model

The electrical model uses an equivalent circuit model (ECM) to describe the battery's electrical behavior through a combination of resistors and capacitors [27]. For DRL-based charging control, a high-precision battery model is established to describe the battery's dynamic characteristics. The model includes a second-order RC model and a thermal model to capture both the electrical and thermal characteristics of LIBs. The voltage dynamics of the RC network are described by the following equations:

$$\frac{dV_{p1}(t)}{dt} = -\frac{V_{p1}(t)}{R_{p1}(t)C_{p1}(t)} + \frac{I(t)}{C_{p1}(t)} \quad (1)$$

$$\frac{dV_{p2}(t)}{dt} = -\frac{V_{p2}(t)}{R_{p2}(t)C_{p2}(t)} + \frac{I(t)}{C_{p2}(t)} \quad (2)$$

$$V_i(t) = V_{oc}(SOC_i(t), t) + V_{p1}(t) + V_{p2}(t) + R_s(t)I(t) \quad (3)$$

where $I(t)$ is the load current, $V_i(t)$ is the terminal voltage of the i th cell at time t , V_{oc} is the open-circuit voltage, $C_{p1}(t)$ is the nominal capacity, and $V_{p1}(t)$ and $V_{p2}(t)$ are the polarization voltages of the two RC circuits.

The SOC of the battery over time can be calculated as:

$$\frac{dSOC_i(t)}{dt} = \frac{I(t)}{3600C_n}, \quad I(t) = a(t) \times C \quad (4)$$

where $SOC_i(t)$ represents the state of charge of the i th cell at time t , and $I(t)$ denotes the charging current at time t . $a(t)$ is the normalized charge rate with respect to time t , introduced to facilitate unified optimization. C_n denotes the nominal capacity of the battery.

2.2. Thermal Model

During charging and discharging, batteries generate heat, primarily due to internal resistance (ohmic heat), polarization processes, and entropy effects from electrochemical reactions [28]. To describe these effects, a lumped-parameter thermal model is used, simplifying the heat conduction problem within the battery by dividing it into core and surface parts.

The core temperature change is described by:

$$\frac{dT_c(t)}{dt} = \frac{1}{C_c} \left(Q_{gen} - \frac{T_c(t) - T_s(t)}{R_{cs}} \right) \quad (5)$$

where $T_c(t)$ is the core temperature at time t , $T_s(t)$ is the surface temperature at time t , C_c is the thermal capacitance of the core, and R_{cs} is the thermal surface-to-coolant thermal barrier.

$$\frac{dT_s(t)}{dt} = \frac{1}{C_s} \left(\frac{T_c(t) - T_s(t)}{R_{cs}} - \frac{T_s(t) - T_f(t)}{R_{cf}} \right) \quad (6)$$

where $T_s(t)$ is the coolant temperature at time t , C_s is the thermal capacitance of the surface, and R_{cf} is the thermal resistance between the surface and coolant.

The heat generation in lithium-ion batteries during operation primarily stems from three fundamental mechanisms: (1) Joule heating (or ohmic heating), caused by internal resistance; (2) entropic heating, resulting from entropy changes during battery reactions; and (3) polarization heating, arising from electrochemical reactions in the battery.

The total heat generated during charging and discharging Q_{gen} consists of three parts:

$$Q_{\text{gen}} = I(t)^2 \cdot R_s + V_p \cdot I(t) + E_n \cdot \Delta T \quad (7)$$

2.3. Lithium-Ion Battery Pack Environment

A new simulation environment for lithium-ion battery packs, compatible with DRL, has been created to model the dynamic behavior of battery packs [28]. The simulation environment incorporates two key elements: a mathematical representation of the battery pack and its corresponding optimization objective, which provide critical feedback for agent training. The simulation environment integrates the aforementioned electrothermal-coupled battery model to establish a standardized platform for developing and evaluating intelligent charging strategies for lithium-ion batteries (LIBs). Built upon the Python Battery Mathematical Modelling (PyBamm) framework (version 24.11.1), this platform incorporates a reinforcement learning interface with enhanced scalability.

Python Battery Mathematical Modelling (PyBamm) is an open-source Python package (version 3.11.4) developed by the Faraday Institute [29], providing reduced-order models for lithium-ion and lead-acid batteries, supporting fast and flexible simulations using ODE/DAE solvers, and allowing modification of existing model parameters or creation of new models and solvers.

Figure 2 shows the simulation results of charging four battery packs with a constant current of 100 A. Figure 2 reveals that under uniform current input conditions, parametric inconsistencies among individual battery cells induce divergent charging rates, consequently leading to significant state of charge (SOC) and terminal voltage imbalances within the battery pack. Notably, these inconsistencies exhibit cumulative effects during cycling, ultimately compromising the overall performance and service life of the battery system.

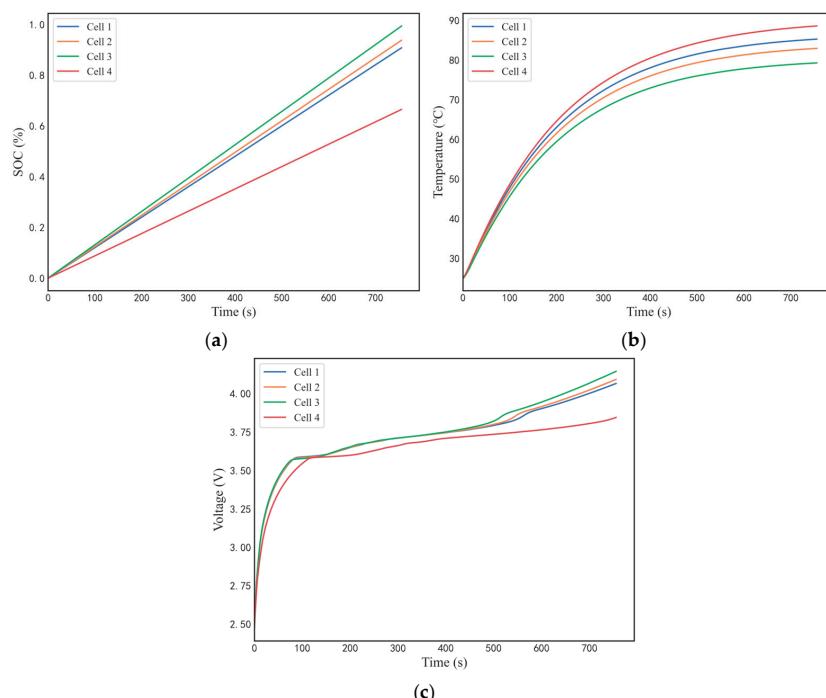


Figure 2. SOC, temperature, and voltage simulation of lithium-ion battery pack charging. (a) SOC simulation results; (b) temperature simulation results; and (c) voltage simulation results.

3. Lithium-Ion Battery Charging Optimization Model

In fast battery charging, achieving an optimal balance between minimizing charging time and maintaining SOC consistency among cells is a key challenge in BMS design.

3.1. Objective Function

To simultaneously optimize charging time and SOC consistency, this paper defines a cost function J , which includes charging time cost C_{SOC} , SOC balancing cost $C_{balance}$, and penalty terms $C_{constraints}$.

$$J = w_1 C_{balance} + w_2 C_{SOC} + w_3 C_{constraints} \quad (8)$$

(1) Charging Time Cost C_{SOC}

C_{SOC} represents the time required to charge from the initial SOC to the target SOC and is defined as:

$$C_{SOC} = e^{\frac{t}{a}} |SOC_i(t) - SOC_{desired}|, \quad SOC(t) = \sum_{i=1}^N SOC_i(t) \quad (9)$$

In the electrical model, $SOC_i(t)$ is the remaining charge of the i th cell at time t , determined by the charging current $a(t)$ and the battery capacity, as given by Equation (4). t denotes the cumulative charging time (in seconds), and a is the time penalty coefficient. The exponential term $e^{\frac{t}{a}}$ acts as a time penalty factor, amplifying the cost of deviation as time increases to encourage faster charging by the agent.

(2) SOC Balancing Cost $C_{balance}$

Manufacturing differences, aging, and temperature gradients often cause SOC variations among the N cells in a battery pack, affecting overall system performance. The SOC balancing cost is introduced to quantify the degree of inconsistency in SOC among the individual cells within the pack. The balancing cost is expressed as:

$$C_{balance} = \sum_{i=1}^N |SOC_i(t) - \overline{SOC}| \quad (10)$$

where $SOC_i(t)$ is the remaining charge of the i th cell at time t , and \overline{SOC} is the average remaining charge of all cells in the battery pack.

(3) Penalty terms $C_{constraints}$

The constraints in the battery charging optimization problem include voltage, thermal, SOC, and state of health (SOH). These constraints ensure the safety and stability of the charging process. The specific expressions are as follows:

$$C_{constraints} = C_{thermal} + C_{voltage} + C_{soh} + C_{SoC} \quad (11)$$

3.2. Constraints

(1) Voltage Constraint

To prevent overcharging, the battery voltage should remain within a safe range. The voltage constraint is expressed as:

$$C_{voltage} = \begin{cases} 0, & \text{if } V_{min} \leq V_i(t) \leq V_{max} \\ |V_i(t) - V_{min}|, & \text{if } V_i(t) < V_{min} \\ |V_i(t) - V_{max}|, & \text{if } V_i(t) > V_{max} \end{cases} \quad (12)$$

where $V_i(t)$ is the voltage of the i th cell at time t , and $V_i(t)$ and $a(t)$ are related to the voltage Equation (3) in the electrical model. The voltage dynamics are influenced by the load current and internal resistance.

(2) Temperature Constraint

Temperature is a critical factor affecting battery safety and performance. During charging, the battery temperature should remain within a safe range to avoid thermal runaway and reduced charging efficiency.

$$C_{\text{thermal}} = \begin{cases} 0, & \text{if } T_{\min} \leq T_i(t) \leq T_{\max} \\ |T_i(t) - T_{\min}|, & \text{if } T_i(t) < T_{\min} \\ |T_i(t) - T_{\max}|, & \text{if } T_i(t) > T_{\max} \end{cases} \quad (13)$$

where $T_i(t)$ is the temperature of the i th cell at time t , and $T_i(t)$ and $a(t)$ are related to the temperature Equations (5) and (6) in the thermal model. Temperature changes are influenced by internal resistance, charging current, and electrochemical reactions.

(3) SOC Constraint

The SOC of the battery must also remain within a safe range to avoid overcharging. The SOC constraint is expressed as:

$$C_{\text{SoC}} = \begin{cases} 0, & \text{if } 0 \leq SOC_i(t) \leq 1 \\ |SOC_i(t) - SOC_{\min}|, & \text{if } SOC_i(t) < 0 \\ |SOC_i(t) - SOC_{\max}|, & \text{if } SOC_i(t) > 1 \end{cases} \quad (14)$$

where $SOC_i(t)$ is the remaining charge of the i th cell at time t , and $a(t)$ is the charging current, as given by Equation (4).

(4) SOH Constraint

The state of health (SOH) reflects the degree of battery performance degradation. During optimization, the SOH should be maintained at a reasonable level to extend battery life and reduce maintenance costs. The SOH constraint is expressed as:

$$C_{\text{soh}} = \begin{cases} 0, & \text{if } 0.7 \leq SOH_i(t) \leq 1 \\ |SOH_i(t) - SOH_{\min}|, & \text{if } SOH_i(t) < 0.7 \\ |SOH_i(t) - SOH_{\max}|, & \text{if } SOH_i(t) > 1 \end{cases} \quad (15)$$

where $SOH_i(t)$ is the SOH of the i th cell at time t , and $SOH_i(t)$ and $a(t)$ are related to Equation (4).

To flexibly handle these constraints, a soft penalty mechanism is adopted, incorporating constraint violations into the cost function through penalty terms $C_{\text{constraints}}$.

3.3. Decision Variables

In this fast-balancing charging optimization problem, the decision variable is the charging current $a(t)$ applied to the battery at each control cycle (t).

The charging current $a(t)$ not only determines the speed of the charging process but also directly affects the SOC balancing, voltage, temperature, and SOH of the battery pack. This paper sets the charging current range to $[0, 6C]$, expressed as:

$$a(t) \in [0, 6C] \quad (16)$$

In this study, the optimization goal of the charging strategy is not a single static value but a dynamic process that continuously adjusts over time t . Common methods to solve such dynamic processes include model predictive control (MPC), reinforcement learning (RL), and deep reinforcement learning (DRL). Among these, the DDPG algorithm, as a DRL method, interacts with the environment to select the charging current $a(t)$ at each time step t as the optimization decision.

This decision not only is based on the current state S_t (SOC, voltage, temperature, etc.) but also considers historical states and future changes, forming a time-varying dynamic optimization process. Specifically, the goal of DRL is to maximize cumulative rewards, which are calculated by weighting multiple objectives (minimizing charging time, SOC consistency, penalty terms). Since the battery state continuously changes during charging, DRL needs to adjust the charging current $a(t)$ at each time step t to ensure the charging process is optimal at every moment, rather than relying on a fixed charging current. This time-dependent dynamic optimization allows the DDPG algorithm to adapt to uncertainties and dynamic changes during the charging process. Each time step's decision not only considers the current battery state but also predicts and responds to future battery state changes, ensuring overall charging efficiency and safety.

3.4. Improved Deep Deterministic Policy Gradient

DDPG, as an actor-critic-based RL algorithm, is particularly suitable for problems with continuous action spaces. Its core idea is to use a deterministic policy (actor) to select actions and evaluate the value of the selected actions through a value function (critic). By combining target networks and experience replay buffers, DDPG effectively improves the stability and efficiency of the training process. The introduction of target networks mitigates the high variance problem of target values, while the experience replay mechanism breaks the correlation between samples, promoting sample diversity and independence. These designs make DDPG perform well in high-dimensional, continuous control tasks. To further enhance the algorithm's stability and robustness, this paper introduces reward centralization and entropy regularization mechanisms, improving the exploration ability and diversity of the policy.

At a specific time step t , DDPG determines the action by considering exploration and the inherent policy, as shown in the following formula:

$$a_t = \mu(s_t | \theta^\mu) + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \sigma^2) \quad (17)$$

In DDPG, the target Q -value is typically expressed as: where s_t is the state space, θ^μ is the parameter of the network μ (decision variable), and ϵ_t is Gaussian noise, which only exists during the training phase.

$$Q^*(s_t, a_t) = r(s_t, a_t) + \gamma Q'(s_{t+1}, \mu'(s_{t+1} | \theta^{\mu'}) | \theta^{Q'}) \quad (18)$$

To improve the stability and learning efficiency of DDPG, this paper introduces a reward centralization mechanism. Reward centralization eliminates constant terms in the value function, allowing the value function to focus more on relative differences between states, improving learning efficiency and stability. By reducing unnecessary fluctuations, the algorithm can converge to the optimal policy more quickly. The constant term is calculated as:

$$\bar{r}_t = \beta \bar{r}_{t-1} + (1 - \beta) \frac{1}{|\mathcal{B}|} \sum_{i=1}^{|\mathcal{B}|} r_i \quad (19)$$

where \bar{r}_t is the sliding mean of the reward at the current time step, β is the sliding average weight factor (set to 0.99), $|\mathcal{B}|$ is the batch size, and r_i is the reward of the i th sample.

Reward centralization adjusts the reward signal by subtracting the mean reward from each time step's reward, centering the reward signal around the mean. The mathematical expression is as follows:

$$r'_i = r_i - \bar{r}_t \quad (20)$$

Specifically, in the network update part of the algorithm, the centralized reward is used to calculate the target Q -value, reducing the variance of the reward signal and improving learning stability.

$$Q^*(s_t, a_t) = r'(s_t, a_t) + \gamma \mathbb{E}_{s_{t+1} \sim \mathcal{E}} [Q'(s_{t+1}, \mu'(s_{t+1} | \theta^{\mu'})) | \theta^{Q'}] \cdot (1 - d_t) \quad (21)$$

where $r'(s_t, a_t)$ is the centralized immediate reward, γ is the discount factor, $Q'(s_{t+1}, a' | \theta^{Q'})$ is the Q -value evaluated by the target Critic network, $\mu'(s_{t+1} | \theta^{\mu'})$ is the action output by the target actor network in state s_{t+1} , and d_t is a binary variable indicating whether the state is terminal.

The Bellman equation shows that the optimal evaluation of the current state and action can be obtained recursively, accurately approximating this iterative task. The critic network update error can be optimized by minimizing the mean squared error (MSE) between the current Q -value and the target Q -value. The loss function is defined as:

$$L^Q(\theta) = -\mathbb{E}_{s \sim \mathcal{D}} [(Q^*(s_t, a_t) - Q(s_t, a_t | \theta^Q))^2] \quad (22)$$

$$a_{t+1} = \mu(s_t | \theta^{\mu}) \quad (23)$$

where $Q(s_t, a_t | \theta^Q)$ is the Q -value evaluated by the current critic network, and $Q^*(s_t, a_t)$ is the target Q -value, calculated using the Bellman equation.

To improve the exploration ability and diversity of the policy, this paper introduces an entropy regularization mechanism. Entropy regularization encourages policy diversity by adding an entropy term to the actor network's loss function, enhancing exploration ability and algorithm robustness. Specifically, the optimization goal of the policy network is to maximize the weighted combination of expected return and policy entropy $\log \pi_\theta(\mu_\theta(s) | s)$. By adjusting the entropy regularization weight parameter α , the policy can pursue high returns while maintaining a certain level of randomness, promoting exploration.

The optimization goal of the policy network is defined as:

$$J(\theta^{\mu}) = \mathbb{E}_{s \sim \mathcal{D}} [Q(s, \mu_\theta(s) | \theta^Q) - \alpha \log \pi_\theta(\mu_\theta(s) | s)] \quad (24)$$

Therefore, the actor network's loss function is optimized by maximizing $J(\theta^{\mu})$, i.e., minimizing the following loss function.

The policy network continuously updates towards the goal of improving performance. To achieve this, the actor network's loss function is defined as the negative gradient of the optimization goal, i.e.:

$$L^{\mu}(\theta) = -\mathbb{E}_{s \sim \mathcal{D}} [Q(s, \mu_\theta(s) | \theta^Q) - \alpha \log \pi_\theta(\mu_\theta(s) | s)] \quad (25)$$

This loss function consists of two parts: the expected return evaluated by $Q(s_t, \mu_\theta(s_t) | \theta^Q)$, aiming to maximize this return, and the entropy term $-\alpha \log \pi_\theta(\mu_\theta(s_t) | s_t)$, encouraging policy diversity and preventing premature convergence to a deterministic solution, thereby improving exploration ability and algorithm robustness.

By minimizing the above loss function L^{μ} , the actor network can optimize the expected return while maintaining policy randomness. This mechanism effectively balances

exploration and exploitation, enhancing the policy's adaptability in complex and dynamic environments.

The introduction of entropy regularization not only encourages policy diversity but also dynamically adjusts the balance between exploration and exploitation, significantly improving the training stability and efficiency of the algorithm. Especially in high-dimensional continuous action spaces, the entropy regularization mechanism provides stronger robustness to the policy, ensuring the algorithm's performance in both training and practical applications.

Updating target networks: The target networks Q' and μ' are updated using a soft update strategy:

$$\theta^{Q'} \leftarrow \tau\theta^Q + (1 - \tau)\theta^{Q'} \quad (26)$$

$$\theta^{\mu'} \leftarrow \tau\theta^{\mu} + (1 - \tau)\theta^{\mu'} \quad (27)$$

4. Fast-Balanced Charging Dynamic Optimization Method

Based on the electrothermal battery model developed in Section 1, this section proposes a deep reinforcement learning (DRL) method using the improved deep deterministic policy gradient (DDPG) algorithm. The method aims to simultaneously optimize the charging time and SOC consistency of lithium-ion batteries (LIBs) to achieve a fast-charging strategy. To formulate the optimal charging strategy, this paper transforms the charging problem into an optimization problem, which includes defining a cost function consistent with the objectives. Specifically, the cost function considers charging time, SOC consistency, and other factors, evaluating the quality of the charging strategy through a weighted combination of these factors. Building upon the modeling framework, this study formulates the lithium-ion battery charging optimization problem as a deep reinforcement learning (DRL) task. The state space includes key parameters such as battery SOC, voltage, and temperature, while the action space covers the charging current at each time step. The specific implementation of each step is discussed in detail below.

4.1. Reward Function

The goal of solving the fast-balance charging problem is to minimize the cost function, as stated in (8). Therefore, to transform the optimal control problem to a DRL problem, we define the reward function as the negative of the cost. Assuming the goal is to minimize the cost function J , the reward function is defined as:

$$r(st, at) = -J_t \quad (28)$$

where $r(st, at)$ represents the reward obtained from the environment by taking action at in state st .

4.2. State Space and Action Space

The state space is typically composed of the system's key state variables. In the charging control task of the BMS, the state space can be defined as the battery's state (SOC), voltage, and temperature. In this paper, the state space is defined as:

$$S = \{SoC, Tc, Vt\} \quad (29)$$

where SoC is the battery's state of charge, representing the proportion of energy currently stored, typically ranging from 0% to 100%, as detailed in Equation (4). Tc is the battery's temperature, as detailed in Equation (7). Vt is the battery's voltage level, as detailed in Equation (7).

The DRL agent should determine the optimal charging current C-rate and SOC consistency to minimize the total charging time and SOC inconsistency while adhering to constraints. Therefore, the DDPG-DRL strategy aims to control the charging current in a continuous manner, and the action space is defined as:

$$A = \{at \mid t \in (0, 6C)\} \quad (30)$$

where the upper limit of 6C is determined based on the specifications of the studied lithium-ion battery (LIB). at represents the charging current at time t , ranging from 0 to 6C, ensuring the charging process remains within safe limits.

This study proposes a DDPG-DRL fast-charging strategy framework for fast-balanced charging of lithium-ion battery packs. The proposed framework is based on the deep deterministic policy gradient (DDPG) algorithm, incorporating a centralized reward function and constructing a state space that captures multidimensional features of the battery pack along with a continuous action space. It effectively enhances charging speed while achieving optimal dynamic balancing of the battery pack. The improved DDPG algorithm proposed in this study interacts with the electrothermal model through a closed-loop dynamic optimization process. This process consists of four primary components: state observation, action decision, environmental feedback, and policy update, allowing the model to adaptively respond to temperature, voltage, and SOC changes during fast charging. The schematic diagram of the DDPG-DRL strategy is shown in Figure 3.

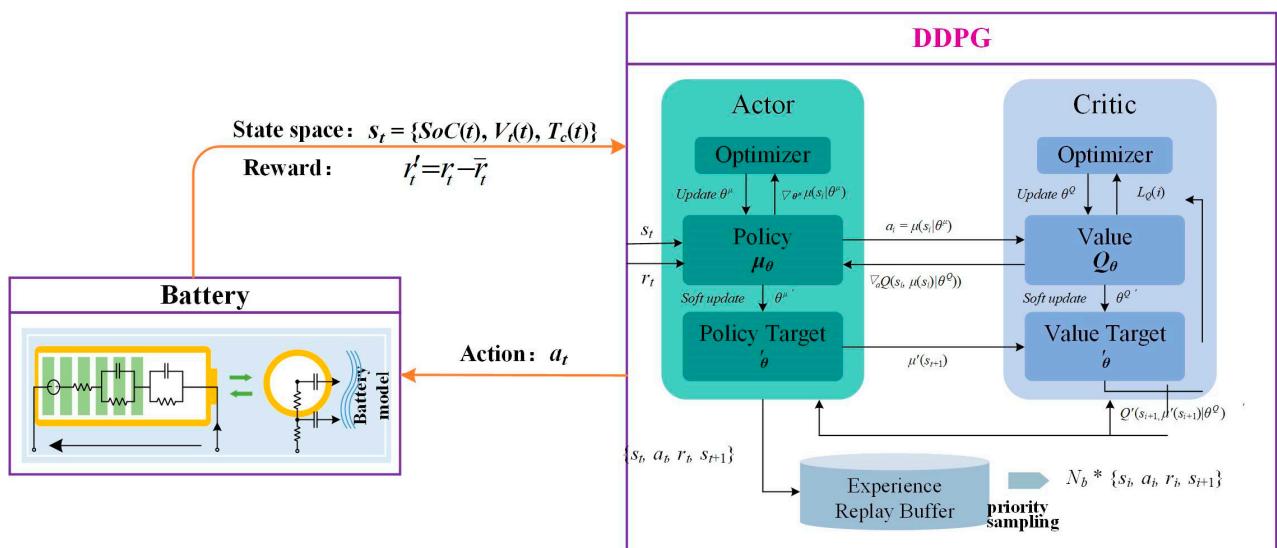


Figure 3. Fast-balanced charging strategy based on improved DDPG.

5. Simulation Analysis

5.1. Battery Model Validation

This study evaluates the accuracy of the established battery model by charging and discharging the A123 commercial LFP battery (ANR23650 M1B) [30]. Specifically, data including input current and output voltage were collected when the environmental temperature was 25 °C. The results show that the electrical and thermal modeling results are very close to the real data under all cycling conditions. As shown in Figure 4, the mean absolute error (MAE) of the modeled voltage is as low as 0.0168. Clearly, the established battery model can accurately describe the electrothermal model of LIBs.

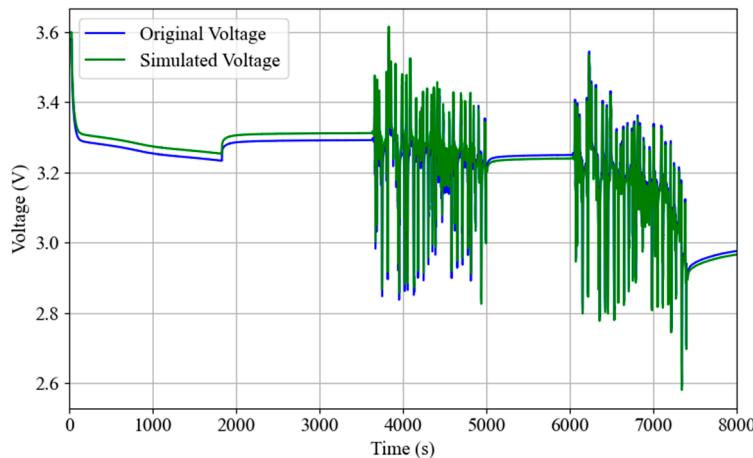


Figure 4. Battery simulation vs. real data voltage comparison.

5.2. Deep Reinforcement Learning Training

In this study, the improved deep deterministic policy gradient (DDPG) model was trained by considering the battery's current state (SOC, voltage, temperature, etc.) at each time step. Through training, the DRL learned how to select a charging current $a(t)$ based on the current state (state space) and historical experience (through the reward function). This dynamic optimization process means that the DRL continuously adjusts its decisions over time t to maximize the final cumulative reward. Therefore, the decision variables (such as charging current) during the charging process are time-varying and continuously optimized throughout the charging process.

From the simulation results, it can be observed that the optimized charging strategy generates a decision curve that varies over time t , rather than a static value. Since the state during the charging process (SOC, temperature, voltage, etc.) continuously changes, the DRL must adjust the charging current $a(t)$ at each time step t based on the current battery state. This dynamic optimization process results in a time-varying curve rather than a single optimized value. This dynamic decision-making process helps achieve higher charging efficiency and safety.

In this study, the improved deep deterministic policy gradient (DDPG) model achieved convergence after approximately 7000 episodes, showing more stable convergence and smaller fluctuations, while achieving a higher average reward throughout the training process. This result indicates that the DDPG model performs well in this task, especially in avoiding overestimation issues, providing more effective exploration opportunities for DRL.

Figure 5 shows the episode return values during the training process. As anticipated, the episode return demonstrates a steady upward trend, reflecting the model's progressive convergence toward an optimal charging strategy, and its performance steadily improves as training progresses.

Figure 6 shows whether the charging time is effectively reduced during DRL training. By optimizing the decision of the charging current $a(t)$, the DRL can more precisely control the charging process, avoiding overcharging and overheating issues, thereby reducing charging time while ensuring safety. The optimized strategy typically increases the charging current when the battery state is suitable, accelerating the charging process. To better show the range of return value changes during training, the 90% confidence interval is shaded in the figure, making the fluctuation range of the return value clearer and further verifying the stability and consistency of the model's performance.

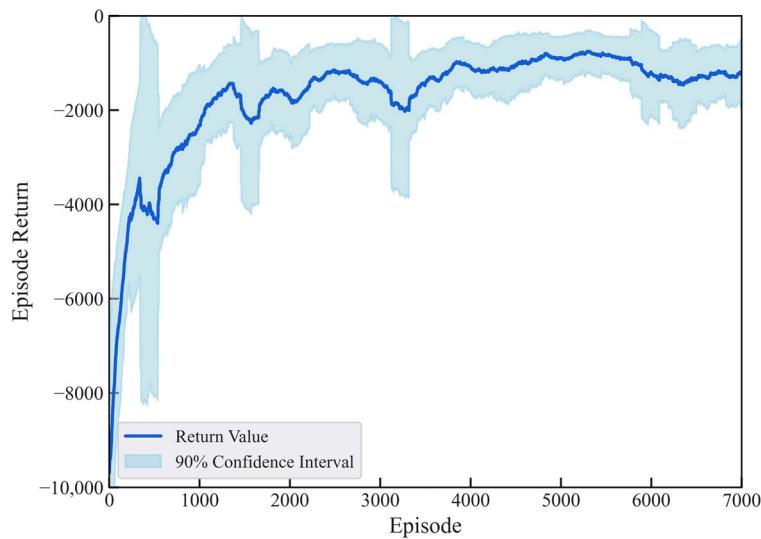


Figure 5. Return value on DRL training.

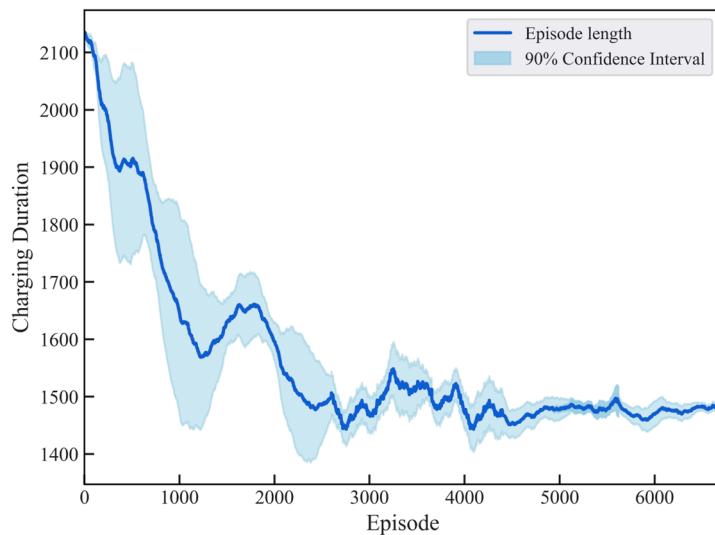


Figure 6. Charging duration for DRL training.

According to the training results, the decision of the charging current shows a time-varying curve, which is in sharp contrast to traditional static charging strategies. The dynamic adjustment ability of DRL makes the charging process more flexible and adaptable to changes in the battery state, avoiding potential risks such as overcharging and overheating, thereby effectively improving charging efficiency and safety.

These training results indicate that the improved DDPG model shows great potential in battery charging optimization problems, especially in achieving dynamic decision-making and efficient optimization.

(1) Hyperparameter Settings

Table 1 details the hyperparameters used during the training of the DDPG model, including their values and descriptions. The reasonable setting of these hyperparameters lays the foundation for efficient training and performance improvement of the model. Specifically, the learning rate is set to 0.001, ensuring that the model parameter update step size maintains training stability while accelerating convergence.

Table 1. Hyperparameters of the DDPG model.

Hyperparameter	Value	Description
Learning Rate	0.001	Learning rate for updating the policy network
Episodes	12,000	Controls the duration of the entire training process
Buffer Size	10,000,000	Stores state transitions experienced by DRL in the environment
Batch Size	64	Batch size of samples used in each training iteration
Discount Factor	0.99	Discount factor for calculating the present value of future rewards
Target Network Soft Update Coefficient	0.005	Soft update coefficient for target network parameter updates
Regularization Coefficient	0.0002	Regularization coefficient for policy and value networks
Time Step (seconds)	1	Time step
Training Rounds	10	Number of training rounds executed in each update cycle

During training, 12,000 episodes were executed, and a buffer size of 10,000,000 was used to store transitions, ensuring the diversity of the experience replay pool and the model's generalization ability. Additionally, a batch size of 256 was used in each training iteration, helping to improve training stability and efficiency.

However, the uncertainties in the model training process primarily stem from overfitting phenomena. Particularly when employing complex network architectures, the bias in the optimization process may lead to excessive sensitivity of the model to specific features, thereby compromising accurate modeling of battery charging dynamics. Therefore, this paper seeks a balance between hyperparameter selection and model complexity to ensure that the model can effectively optimize the charging strategy while maintaining good generalization performance.

(2) Weight Coefficient Settings

In practical battery charging optimization problems, due to the conflicting objectives of charging time, SOC consistency, temperature and voltage control, and battery health, the initial weight coefficients are set based on prior knowledge and expert experience. Charging time is the most important optimization objective in lithium battery charging, so the weight coefficient for charging time is larger, while the weights for other objectives are adjusted based on the actual distribution of optimization objectives to achieve a balance between objectives. To further optimize the balance between objectives, this paper uses sensitivity analysis to evaluate the impact of different weight configurations on optimization results and performs single-variable factor analysis on each weight coefficient. By observing the changes in the objective function value, the weight configuration with the least impact on the objective function and strong stability is ultimately selected.

The results of the sensitivity analysis are shown in Figure 7. The evolution of the cost function demonstrates a two-phase optimization characteristic. During the initial phase, it exhibits a quasi-linear trend strongly correlated with minimal charging duration, indicating the system's prioritization of charging rate optimization in its control strategy. In this phase, regardless of secondary objectives, the optimization process always prioritizes reducing charging time. After reaching a certain limit, gradient of the cost function changes, and the optimization objective gradually shifts to SOC balancing cost. At the third weight combination, the cost function shows a clear inflection point. After this point, further reducing the weight coefficient corresponding to charging time does not significantly change the SOC balancing cost and constraint attenuation. Therefore, the third weight combination can better balance the three optimization objectives of charging time, SOC balancing cost, and constraints. Based on this analysis, the system achieves optimal performance under the following weights: $w_1 = 0.540$, $w_2 = 0.340$, and $w_3 = 0.120$.

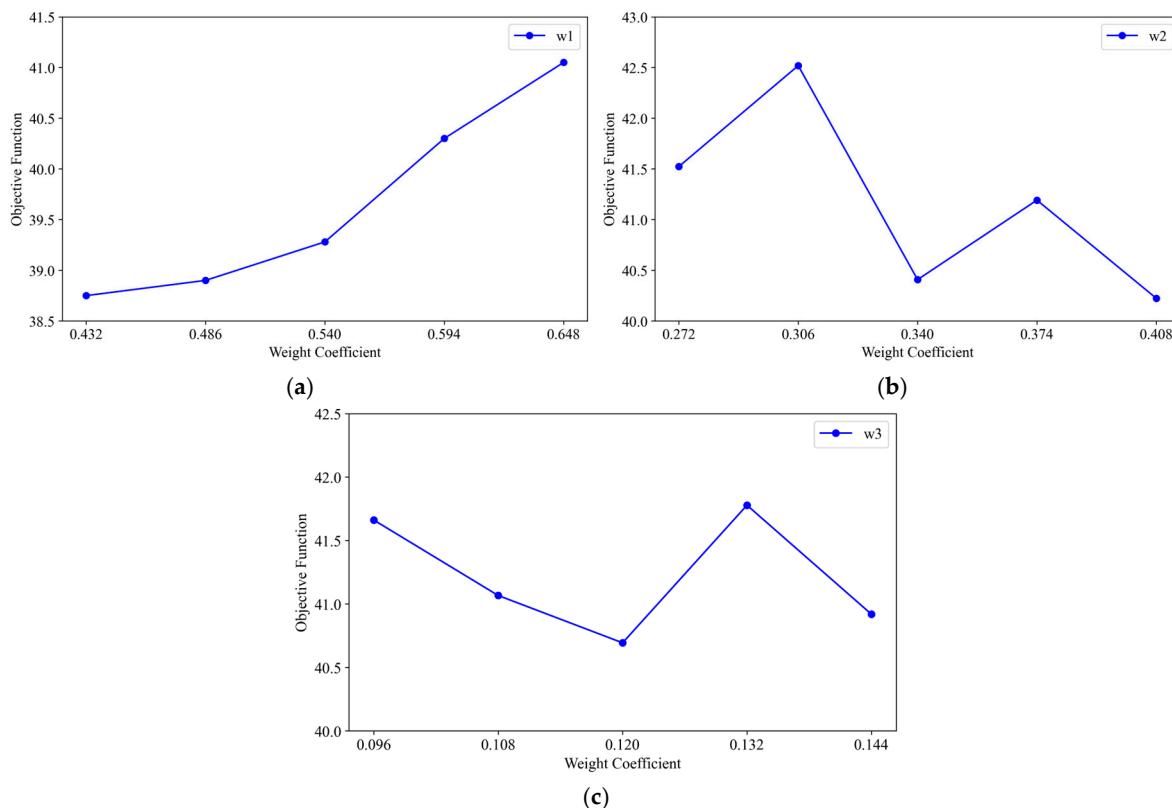


Figure 7. Weight coefficient change curve. (a) Weight coefficient w1; (b) weight coefficient w2; (c) and weight coefficient w3.

5.3. Simulation Results

This section discusses the simulation results of the proposed method, showing the algorithm's performance in a battery pack environment composed of four cells. Each cell's parameters follow a certain Gaussian distribution, showing some differences. The proposed charging strategy first ensures that all cells reach a consistent SOC through a balancing step to eliminate initial imbalances. Then, the algorithm applies the maximum current for efficient charging while adhering to preset constraints, achieving a fast and safe battery charging process [31,32].

This paper selects a typical case with a low initial SOC and small inconsistency and compares the proposed charging strategy with two other methods (the CC-CV-based charging algorithm and the basic DDPG-based charging strategy) [33,34]. The rule-based charging algorithm integrates passive balancing within the CC-CV charging approach and is utilized in the battery pack charging process. In contrast, the model-based charging method is designed by introducing the basic DDPG algorithm and is compared with the improved method proposed in this study. To ensure a fair comparison, the validation environment for all strategies remains consistent.

From Figure 8a–c, it can be observed that the proposed improved DDPG algorithm outperforms other methods in multiple performance metrics (such as full charging time and balancing time). Figure 8b shows that the traditional CC-CV method with a balancing mechanism has the following limitations: When the highest voltage of the battery pack reaches the set cutoff threshold, the charging mode switches from constant current (CC) to constant voltage (CV). At this time, charging control and balancing control are executed separately, which may lead to the following potential problems: (1) Since the charging current significantly decreases in CV mode, the battery may not be fully charged, prolonging the full charging time, and (2) during the CC charging phase, the large charging current

may exacerbate the SOC inconsistency of the battery pack, further prolonging the balancing time [35–38]. From the comparison of the simulation results in Figure 8a, it is clear that the proposed improved DDPG algorithm overcomes these defects by combining balancing control and fast-charging control, improving overall charging efficiency and effectiveness.

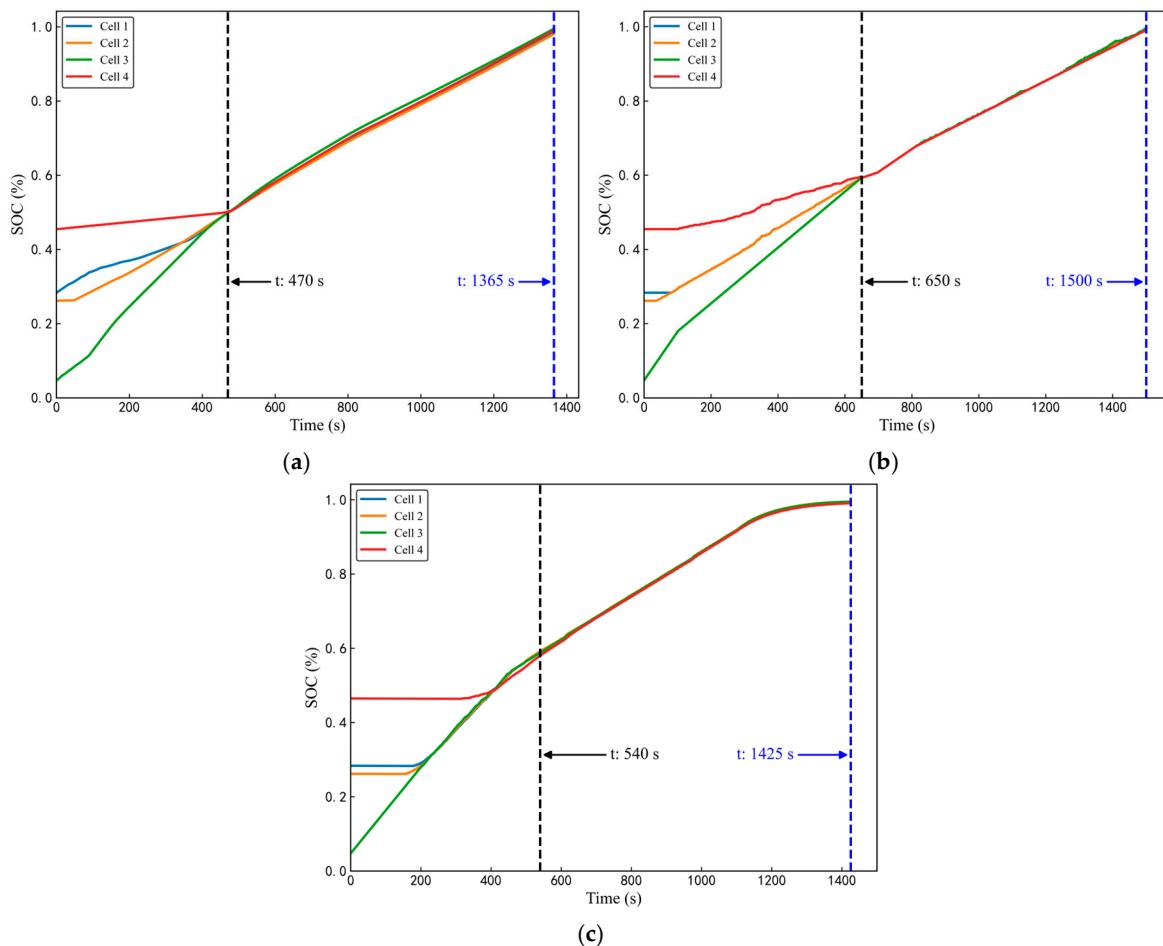


Figure 8. SOC over time based on improved DDPG, base DDPG, and CC-CV strategies for four cells. (a) Improved DDPG; (b) DDPG-based; and (c) CC-CV strategies.

Specifically, the total charging time of the CC-CV method with balancing is 1425 s, while the proposed improved DDPG algorithm only takes 1365 s, saving 60 s of charging time. Additionally, the proposed algorithm can charge more energy with higher efficiency. By analyzing the balancing time during charging, the CC-CV method with balancing requires 540 s for the balancing phase, while the proposed improved DDPG algorithm only needs 470 s. In contrast, the basic DDPG method requires 650 s, showing a clear efficiency advantage.

In summary, the proposed improved DDPG algorithm not only effectively reduces charging time but also significantly reduces battery pack inconsistency, improving overall charging efficiency, verifying its feasibility and superiority in practical applications.

During the charging process, voltage and temperature are important factors affecting battery performance and safety. To more comprehensively evaluate the performance of the proposed basic DDPG strategy, this section further compares and analyzes the voltage change characteristics under three charging strategies. Figure 9 shows the battery voltage curves under different charging strategies, with specific analysis as follows.

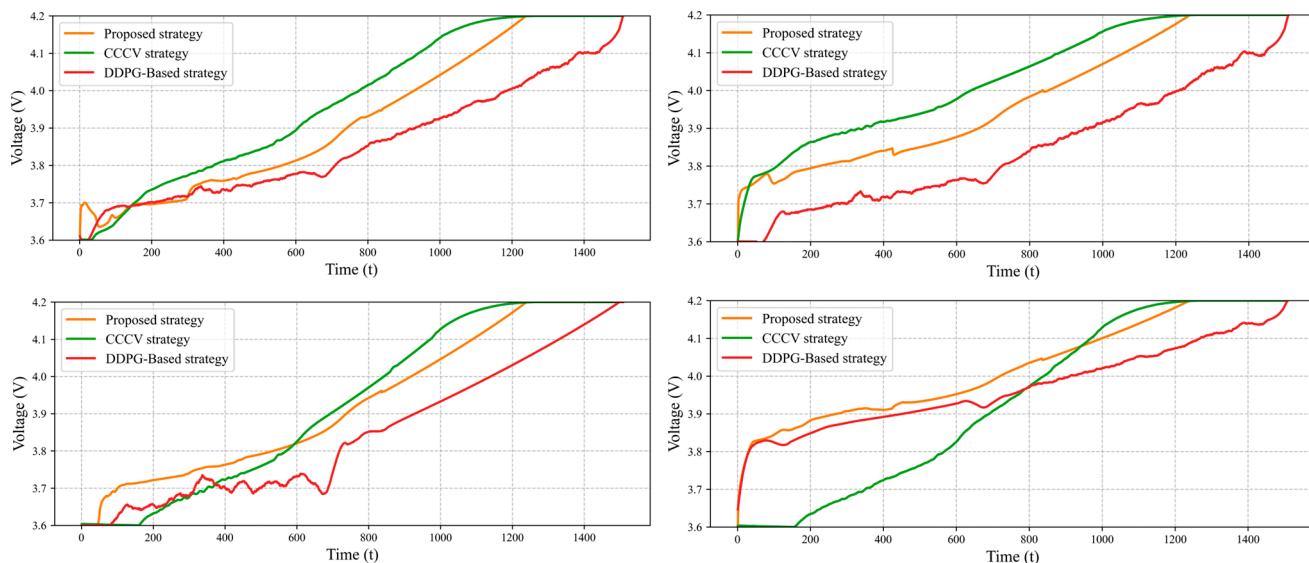


Figure 9. Comparison of terminal voltages under different strategies for four cells.

First, In the CC-CV strategy, during the constant current (CC) phase, high currents cause the battery voltage to rise rapidly, prematurely reaching the maximum voltage threshold and forcing the charging process into the constant voltage (CV) phase, which reduces the overall charging speed. As the battery approaches full charge, the rapid voltage increase may exceed the safe range, triggering overvoltage issues that threaten battery safety and decrease charging efficiency. Although the CV phase prevents overvoltage by lowering the current, in fast-charging scenarios, local high ion concentrations and polarization can still induce lithium plating during the high-current CC phase. The proposed DDPG method mitigates this risk by dynamically adjusting the current for each battery cell and enforcing thermal constraints, thereby reducing the likelihood of lithium plating and subsequent capacity degradation.

In contrast, the proposed improved DDPG strategy intelligently adjusts the charging current to achieve precise control of the battery voltage. In the early stages of charging, the battery voltage rises slowly, and the charging current gradually increases to ensure that the battery remains within a safe operating range during the charging process. Since the basic DDPG strategy fully considers battery thermal management and voltage control, the battery voltage is always maintained within the defined safe range, avoiding overvoltage phenomena and ensuring a healthy charging process. This control strategy not only helps extend the battery's lifespan but also minimizes losses caused by voltage fluctuations.

Additionally, the voltage curves in Figure 9 show the voltage changes under the third strategy (i.e., the basic DDPG-based charging strategy). Although this strategy can quickly charge the battery to the target SOC in some cases, the voltage fluctuations are larger, and the charging time is longer.

From the comparison results in Figure 8, the proposed basic DDPG strategy takes 650 s to charge the lithium-ion battery to the target SOC, which is 110 s longer than the benchmark strategy. This conservative charging method is driven by stringent limits on the charging current to ensure the battery's thermal performance is maintained and to prevent degradation during charging. As shown in Figure 10, the CC-CV protocol shows a clear trade-off between charging speed and constraint compliance. Specifically, although the CC-CV strategy significantly reduces the overall charging time, its fast-charging process causes the battery core temperature to rise rapidly, exceeding the safe threshold in the early stages and reaching about 51 °C. This temperature overshoot not only endangers the safety and lifespan of the lithium-ion battery (LIB) but may also be considered a slight abuse of the battery, reducing its overall reliability. In contrast, the proposed basic DDPG strategy

effectively controls the charging current, keeping the internal temperature of the battery below the defined safety threshold [39,40].

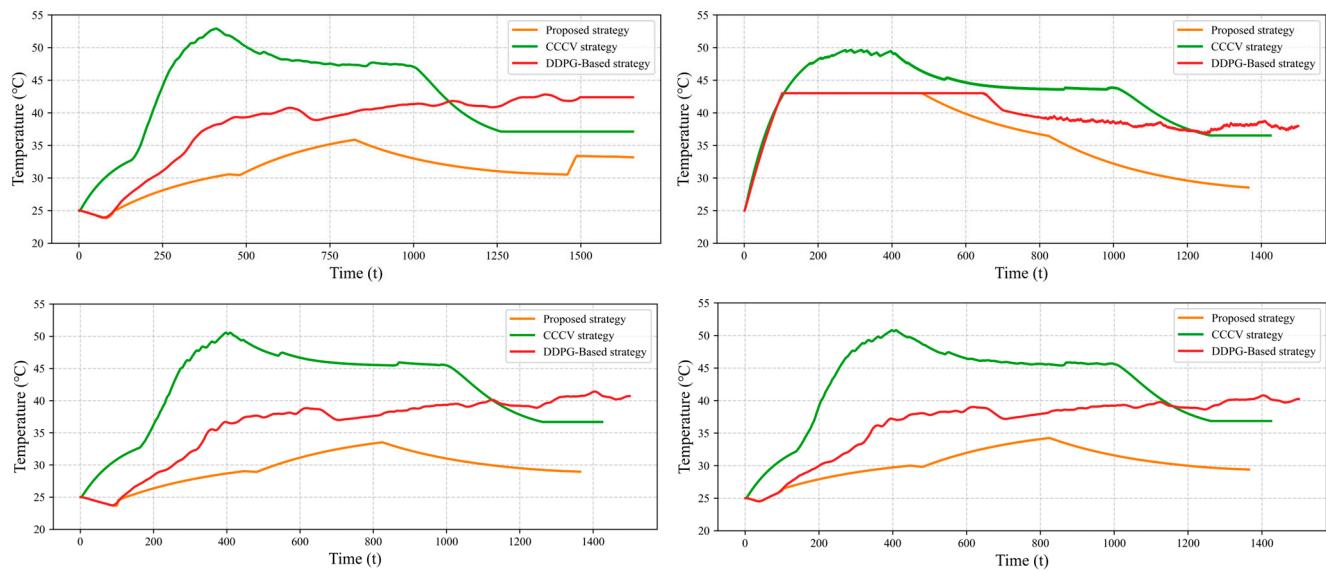


Figure 10. Comparison of temperatures under different strategies for four batteries.

This temperature control effect is consistent with expectations, as high temperatures not only affect the optimization process of the charging strategy by introducing additional “costs” (such as thermal penalties) but may also accelerate the degradation of battery materials, shortening the battery’s lifespan.

However, overly conservative charging strategies, while effectively suppressing temperature rise, inevitably affect charging speed, leading to longer charging times. Table 2 shows the comparison results. In contrast, the improved DDPG algorithm can ensure charging speed while effectively controlling temperature rise, avoiding the unsafe phenomena observed in the CC-CV strategy. This advantage is mainly attributed to the adaptive ability of the DDPG algorithm in the dynamic optimization process, which adjusts the charging strategy based on the real-time state of the battery, avoiding the limitations of fixed current modes. Therefore, the CC-CV charging strategy, due to its arbitrary and fixed current mode, fails to achieve optimal charging optimization, further verifying the superiority of the improved DDPG-based charging strategy in practical applications.

Table 2. Comparison of different charging strategies.

Charging Strategy	Metric	Quantitative Result
Improved DDPG	Balancing Time	470
	Full Charging Time	1365
	Peak Temperature	42
	Peak Voltage	4.2
	Minimum Cycle Cost	1337
DDPG	Balancing Time	650
	Full Charging Time	1500
	Peak Temperature	42.5
	Peak Voltage	4.2
	Minimum Cycle Cost	1593
CC-CV	Balancing Time	540
	Full Charging Time	1425
	Peak Temperature	52
	Peak Voltage	4.2
	Minimum Cycle Cost	15,082

6. Conclusions

This paper addresses the issue of reduced overall performance and available capacity of lithium-ion batteries during fast charging due to cell imbalance by proposing an innovative charging method based on reinforcement learning. By employing the deep deterministic policy gradient (DDPG) algorithm integrated with reward centralization and entropy regularization mechanisms, the charging current is dynamically adjusted to achieve an optimal balance between fast charging and battery health.

Through simulation experiments, the proposed method significantly reduces the total charging time and balancing time compared to the traditional constant-current constant-voltage (CC-CV) strategy, saving 60 s and 70 s, respectively. Additionally, compared to the basic DDPG method, the improved algorithm shows a clear advantage in charging efficiency. This indicates that the proposed method not only effectively improves charging efficiency but also extends battery life and ensures the safety of the charging process.

The main conclusions of this paper are as follows.

- (1) Combining the electrothermal battery model with a continuous deep reinforcement learning algorithm: This enables more accurate simulation of the dynamic performance characteristics of lithium-ion batteries, achieving more efficient charging control.
- (2) Integrating reward centralization and entropy regularization mechanisms into the DDPG algorithm: Reward centralization optimizes the reward function, making the states of battery cells more coordinated, while the entropy regularization mechanism enhances the algorithm's exploration ability and policy diversity by increasing policy randomness, improving training stability and optimization effectiveness.
- (3) Designing a cost function for target optimization: A cost function that comprehensively considers charging time and SOC balancing is constructed, with appropriate penalty terms introduced to effectively achieve target optimization, ensuring battery state consistency while minimizing charging time.

Despite promising results, several limitations remain. First, the proposed method has only been validated in simulation; its effectiveness under real-world conditions—considering model uncertainties, sensor noise, and hardware constraints—requires further study. Second, the computational complexity of the enhanced DDPG algorithm may limit its applicability in embedded systems. Third, the assumption of ideal cell state monitoring may not hold in practical battery management scenarios. Future work will focus on hardware-in-the-loop validation, model simplification for real-time embedded deployment, and improving robustness against uncertainties. These efforts aim to bridge the gap between algorithmic innovation and practical application in intelligent battery charging.

Author Contributions: Conceptualization, X.P.; methodology, Z.Z. (Zhi Zhang) and X.P.; software, Z.Z. (Zhi Zhang); validation, Z.Z. (Zhi Zhang); formal analysis, Z.Z. (Zhi Zhang); investigation, Z.Z. (Zhi Zhang) and X.P.; resources, X.P.; data curation, Z.Z. (Zhi Zhang); writing—original draft, Z.Z. (Zhi Zhang) and X.P.; writing—review and editing, Z.Z. (Zedong Zheng); visualization, Z.Z. (Zhi Zhang) and T.G.; supervision, Y.L.; project administration, Y.L.; funding acquisition, Z.Z. (Zhi Zhang) and X.P. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Natural Science Foundation of Liaoning Province of China (2023JH2/101700068, 2024-MS-217), the Department of Education of Liaoning Province of China (LJ222411632035, LJ212411632075), and the Shenyang Science and Technology Plan Project (24-213-3-29).

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare that they have no conflicts of interest.

References

1. Rahimi-Eichi, H.; Ojha, U.; Baronti, F.; Chow, M.Y. Battery management system: An overview of its application in the smart grid and electric vehicles. *IEEE Ind. Electron. Mag.* **2013**, *7*, 4–16. [[CrossRef](#)]
2. Lin, X.; Khosravinia, K.; Hu, X.; Li, J.; Lu, W. Lithium plating mechanism, detection, and mitigation in lithium-ion batteries. *Prog. Energy Combust. Sci.* **2021**, *87*, 100953. [[CrossRef](#)]
3. Naguib, M.; Kollmeyer, P.; Emadi, A. Lithium-ion battery pack robust state of charge estimation, cell inconsistency, and balancing. *IEEE Access* **2021**, *9*, 50570–50582. [[CrossRef](#)]
4. Tang, X.; Gao, F.; Liu, K.; Liu, Q.; Foley, A.M. A balancing current ratio based state-of-health estimation solution for lithium-ion battery pack. *IEEE Trans. Ind. Electron.* **2021**, *69*, 8055–8065. [[CrossRef](#)]
5. Liao, L.; Chen, H. Research on two-stage equalization strategy based on fuzzy logic control for lithium-ion battery packs. *J. Energy Storage* **2022**, *50*, 104321. [[CrossRef](#)]
6. He, T.; Kang, X.; Wang, F.; Zhang, J.; Zhang, T.; Ran, F. Capacitive contribution matters in facilitating high power battery materials toward fast-charging alkali metal ion batteries. *Mater. Sci. Eng.* **2023**, *154*, 100737. [[CrossRef](#)]
7. Zeng, X.; Li, J.; Qiao, L.; Chen, M. Experimental study on the performance of power battery module heating management under a low temperatures charging scenario. *Int. J. Heat Mass Transf.* **2024**, *225*, 125388. [[CrossRef](#)]
8. Cao, J.; Schofield, N.; Emadi, A. Battery balancing methods: A comprehensive review. In Proceedings of the 2008 IEEE Vehicle Power and Propulsion Conference, Harbin, China, 3–5 September 2008; IEEE: Piscataway, NJ, USA, 2008; pp. 1–6.
9. Xu, J.; Mei, X.; Wang, H.; Shi, H.; Sun, Z.; Zou, Z. A model based balancing system for battery energy storage systems. *J. Energy Storage* **2022**, *49*, 104114. [[CrossRef](#)]
10. Kelkar, A.; Dasari, Y.; Williamson, S.S. A comprehensive review of power electronics enabled active battery cell balancing for smart energy management. In Proceedings of the 2020 IEEE International Conference on Power Electronics, Smart Grid and Renewable Energy (PESGRE2020), Cochin, India, 2–4 January 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1–6.
11. Zhu, Z.; Dong, G.; Lou, Y.; Sun, L.; Yu, J.; Wu, L.; Wei, J. MPC-Guided Deep Reinforcement Learning for Optimal Charging of Lithium-Ion Battery with Uncertainty. *IEEE Trans. Transp. Electrif.* **2024**, *19*, 1128. [[CrossRef](#)]
12. Zhu, R.; Peng, W.; Yang, F.; Xie, M. Fast Charging Protocols Design of Lithium-ion Battery: A Multiple Objective Bayesian Optimization Perspective. *IEEE Trans. Transp. Electrif.* **2025**; *early access*.
13. Khan, A.B.; Choi, W. Optimal charge pattern for the high-performance multistage constant current charge method for the li-ion batteries. *IEEE Trans. Energy Convers.* **2018**, *33*, 1132–1140. [[CrossRef](#)]
14. Zhang, S.S.; Xu, K.; Jow, T.R. Study of the charging process of a LiCoO₂-based Li-ion battery. *J. Power Sources* **2006**, *160*, 1349–1354. [[CrossRef](#)]
15. Shen, Y.; Wang, X.; Jiang, Z.; Luo, B.; Chen, D.; Wei, X.; Dai, H. Online detection of lithium plating onset during constant and multistage constant current fast charging for lithium-ion batteries. *Appl. Energy* **2024**, *370*, 123631. [[CrossRef](#)]
16. Yu, C.; Huang, S.; Xu, H.; Yan, J.; Rong, K.; Sun, M. Optimal charging of lithium-ion batteries based on lithium precipitation suppression. *J. Energy Storage* **2024**, *82*, 110580. [[CrossRef](#)]
17. Liu, D.; Wang, S.; Fan, Y.; Fernandez, C.; Blaabjerg, F. An optimized multi-segment long short-term memory network strategy for power lithium-ion battery state of charge estimation adaptive wide temperatures. *Energy* **2024**, *304*, 132048. [[CrossRef](#)]
18. Andersson, M.; Streb, M.; Prathimala, V.G.; Siddiqui, A.; Lodge, A.; Klass, V.L.; Lindbergh, G. Electrochemical model-based aging-adaptive fast charging of automotive lithium-ion cells. *Appl. Energy* **2024**, *372*, 123644. [[CrossRef](#)]
19. Tuchnitz, F.; Ebelt, N.; Schlund, J.; Pruckner, M. Development and evaluation of a smart charging strategy for an electric vehicle fleet based on reinforcement learning. *Appl. Energy* **2021**, *285*, 116382. [[CrossRef](#)]
20. Park, S.; Pozzi, A.; Whitmeyer, M.; Perez, H.; Kandel, A.; Kim, G.; Moura, S. A deep reinforcement learning framework for fast charging of Li-ion batteries. *IEEE Trans. Transp. Electrif.* **2022**, *8*, 2770–2784. [[CrossRef](#)]
21. Miao, Y.; Gao, Y.; Liu, X.; Liang, Y.; Liu, L. Analysis of State-of-Charge Estimation Methods for Li-Ion Batteries Considering Wide Temperature Range. *Energies* **2025**, *18*, 1188. [[CrossRef](#)]
22. Zhang, J.; Jing, W.; Lu, Z.; Wu, H.; Wen, X. Collaborative strategy for electric vehicle charging scheduling and route planning. *IET Smart Grid* **2024**, *7*, 628–642. [[CrossRef](#)]
23. Guo, X.; Peng, J.; He, H.; Wu, C.; Zhang, H.; Ma, C. Integrated thermal-energy management for electric vehicles in high-temperature conditions using hierarchical reinforcement learning. *Expert Syst. Appl.* **2025**, *276*, 127221. [[CrossRef](#)]
24. Hao, Y.; Lu, Q.; Wang, X.; Jiang, B. Adaptive model-based reinforcement learning for fast-charging optimization of lithium-ion batteries. *IEEE Trans. Ind. Inform.* **2023**, *20*, 127–137. [[CrossRef](#)]
25. Wang, K.; Wang, H.; Yang, Z.; Feng, J.; Li, Y.; Yang, J.; Chen, Z. A transfer learning method for electric vehicles charging strategy based on deep reinforcement learning. *Appl. Energy* **2023**, *343*, 121186. [[CrossRef](#)]

26. Zhou, H.; He, Q.; Li, Y.; Wang, Y.; Wang, D.; Xie, Y. Enhanced Second-Order RC Equivalent Circuit Model with Hybrid Offline–Online Parameter Identification for Accurate SoC Estimation in Electric Vehicles under Varying Temperature Conditions. *Energies* **2024**, *17*, 4397. [[CrossRef](#)]
27. Zhu, Z.Q.; Liang, D. Perspective of thermal analysis and management for permanent magnet machines, with particular reference to hotspot temperatures. *Energies* **2022**, *15*, 8189. [[CrossRef](#)]
28. Mertin, G.K.; Wycisk, D.; Oldenburger, M.; Stoye, G.; Fill, A.; Birke, K.P.; Wieck, A.D. Dynamic measurement of the entropy coefficient for battery cells. *J. Energy Storage* **2022**, *51*, 104361. [[CrossRef](#)]
29. Sultanuddin, S.J.; Vibin, R.; Kumar, A.R.; Behera, N.R.; Pasha, M.J.; Baseer, K.K. Development of improved reinforcement learning smart charging strategy for electric vehicle fleet. *J. Energy Storage* **2023**, *64*, 106987. [[CrossRef](#)]
30. Sulzer, V.; Marquis, S.G.; Timms, R.; Robinson, M.; Chapman, S.J. Python battery mathematical modelling (PyBaMM). *J. Open Res. Softw.* **2021**, *9*, 14. [[CrossRef](#)]
31. Hou, Y.; Liu, L.; Wei, Q.; Xu, X.; Chen, C. A novel DDPG method with prioritized experience replay. In Proceedings of the 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Banff, AB, Canada, 5–8 October 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 316–321.
32. Vartanian, C.; Bentley, N. A123 systems’ advanced battery energy storage for renewable integration. In Proceedings of the 2011 IEEE/PES Power Systems Conference and Exposition, Phoenix, AZ, USA, 20–23 March 2011; IEEE: Piscataway, NJ, USA, 2011; pp. 1–6.
33. Tan, J.; Chen, X.; Bu, Y.; Wang, F.; Wang, J.; Huang, X.; Zhao, Y. Incorporating FFTA based safety assessment of lithium-ion battery energy storage systems in multi-objective optimization for integrated energy systems. *Appl. Energy* **2024**, *367*, 123472. [[CrossRef](#)]
34. Wu, X.; Liao, B.; Su, Y.; Li, S. Multi-objective and multi-algorithm operation optimization of integrated energy system considering ground source energy and solar energy. *Int. J. Electr. Power Energy Syst.* **2023**, *144*, 108529. [[CrossRef](#)]
35. Jaguemont, J.; Bardé, F. A critical review of lithium-ion battery safety testing and standards. *Appl. Therm. Eng.* **2023**, *231*, 121014. [[CrossRef](#)]
36. Liu, H.; Naqvi, I.H.; Li, F.; Liu, C.; Shafiei, N.; Li, Y.; Pecht, M. An analytical model for the CC-CV charge of Li-ion batteries with application to degradation analysis. *J. Energy Storage* **2020**, *29*, 101342. [[CrossRef](#)]
37. Wu, X.; Hu, C.; Du, J.; Sun, J. Multistage CC-CV charge method for Li-ion battery. *Math. Probl. Eng.* **2015**, *2015*, 294793. [[CrossRef](#)]
38. Erdogan, B.; Savrun, M.M.; Körögölu, T.; Cuma, M.U.; Tümay, M. An improved and fast balancing algorithm for electric heavy commercial vehicles. *J. Energy Storage* **2021**, *38*, 102522. [[CrossRef](#)]
39. Liao, L.; Li, H.; Li, H.; Jiang, J.; Wu, T. Research on equalization scheme of lithium-ion battery packs based on consistency control strategy. *J. Energy Storage* **2023**, *73*, 109193. [[CrossRef](#)]
40. Jiaqiang, E.; Zhang, B.; Zeng, Y.; Wen, M.; Wei, K.; Huang, Z.; Deng, Y. Effects analysis on active equalization control of lithium-ion batteries based on intelligent estimation of the state-of-charge. *Energy* **2022**, *238*, 121822.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.