2023-01-1681

Reinforcement Learning Based Fast Charging of Electric Vehicle Battery Packs

Mohammad Hossein Abbasi[a*], Ziba Arjmandzadeh[b], Jiangfeng Zhang[a*], Bin Xu[b], and Venkat Krovi[a]
[a:] Department of Automotive Engineering, Clemson University, Greenville, SC 29607, USA
[b:] Department of Aerospace and Mechanical Engineering, the University of Oklahoma, Norman, OK 73019, USA

Abstract

Range anxiety and lack of adequate access to fast charging are proving to be important impediments to electric vehicle (EV) adoption. While many techniques to fast charging EV batteries (model-based & model-free) have been developed, they have focused on a single Lithium-ion cell. Extensions to battery packs are scarce, often considering simplified architectures (e.g., series-connected) for ease of modeling. Computational considerations have also restricted fast-charging simulations to small battery packs, e.g., four cells (for both series and parallel connected cells). Hence, in this paper, we pursue a model-free approach based on reinforcement learning (RL) to fast charge a large battery pack (comprising 444 cells). Each cell is characterized by an equivalent circuit model coupled with a second-order lumped thermal model to simulate the battery behavior. After training the underlying RL, the developed model will be straightforward to implement with low computational complexity. In detail, we utilize a Proximal Policy Optimization (PPO) deep RL as the training algorithm. The RL is trained in such a way that the capacity loss due to fast charging is minimized. The pack's highest cell surface temperature is considered an RL state, along with the pack's state of charge. Finally, in a detailed case study, the results are compared with the constant current-constant voltage (CC-CV) approach, and the outperformance of the RL-based approach is demonstrated. Our proposed PPO model charges the battery as fast as a CC-CV with a 5C constant stage while maintaining the temperature as low as a CC-CV with a 4C constant stage.

Introduction

Powering electric vehicles (EVs) with renewable energy is the key solution to mitigate our carbon footprint, 28% of which is attributed to the transportation sector [1,2]. However, the main barrier against EV development is its charging speed, which is much slower than a conventional vehicle's refueling time [3]. Moreover, EVs struggle with battery heating during fast charging, which can lead to thermal runaway. Hence, battery temperature should be kept between 20-40℃ [4]. Consequently, numerous studies were carried out on optimizing battery fast charging. But the majority of the literature either focuses on cell level research or studies small battery packs. There is a lack of study on the execution of fast charging optimization results on a large battery pack. Accordingly, this paper proposes an approach based on deep reinforcement learning (RL) to fast charge a battery pack consisting of 444 lithium-ion cells.

In Ref. [5], an electro-thermal-aging model is developed to charge a battery pack. However, the model assumes identical cells, where cell aging and temperature differences are ignored. The study in [6] utilizes an electrochemical Doyle–Fuller–Newman model coupled with a thermal model to describe the battery, but the study is limited to one cell. Similarly, authors in [7] use an electrochemical model enhanced with a thermal and a degradation model to fast charge a battery pack using model predictive control (MPC). However, in the case study, they only used a battery pack with four cells and did not discuss how fast their proposed model could solve the optimization problem. The research in [8] exploits neural networks to find the optimal charging current using a set of pre-determined current values. However, the work in [8] is limited in three ways. First, it only considers five different values for charging the current C-rate. Second, the maximum C-rate in the list is 2.5C which is relatively small considering that the extremely fast charging target set by the US Department of Energy is 6C or more [9]. Third, the case study considers eight cells connected in series, which has only a few cells with no parallel connection.

Model-free RL can be used to solve optimization problems by casting the problems into a Markov decision process format. A reward is used to train the RL agent to take the optimal actions, which would lead to the maximum cumulative reward. Since we center our work around a model-free RL in this work, our proposed model is independent of the system dynamics, and in turn, it is robust to system and model uncertainties. Ref. [10] demonstrates charging a single battery cell with the proposed RL algorithm outperforms MPC and CC-CV. However, the study is restricted to a cell and is not extended to the pack level. Likewise, Ref. [11] conducts research on a single lithium-ion cell that is optimally fast charged through an adaptive model-based RL algorithm. In addition, the study in [12] explores a balancing-aware fast charging of a battery pack using deep Q-network (DQN). Although the authors formulate the problem for a battery pack, the underlying case study only considers a pack of four cells. Moreover, the work in [12] employs DQN, which can only handle discrete action spaces. Hence, a finite set of action choices are available for the fast-charging algorithm, which differs from real-world problems with continuous action spaces.

There is a lack of studies where a large battery pack is used in developing, simulating, or testing fast charging optimization algorithms. Thus, in this paper, we propose a deep RL approach to optimally fast charge a large battery pack. The said deep RL is trained through proximal policy optimization (PPO). PPO is selected for its simple implementation and fast training compared to some

other approaches, such as deep deterministic policy gradient. Finally, the contributions of this paper are as follows: i) A model-free deep RL is developed to fast charge a battery pack consisting of 444 cells, ii) Each cell in the pack is modeled by an electro-thermal model taking into consideration the temperature gradient among the cells and the discrepancy in the currents of parallel cells, and iii) In the simulations, we demonstrate the benefits of our model over CC-CV.

Modeling

This section describes the equivalent circuit model, the thermal model, and the aging model that are used in this paper.

Battery model

As shown in Figure 1, the battery is represented by a second-order equivalent circuit model comprising a voltage source, two resistor-capacitor (RC) pairs, and a connection resistor.

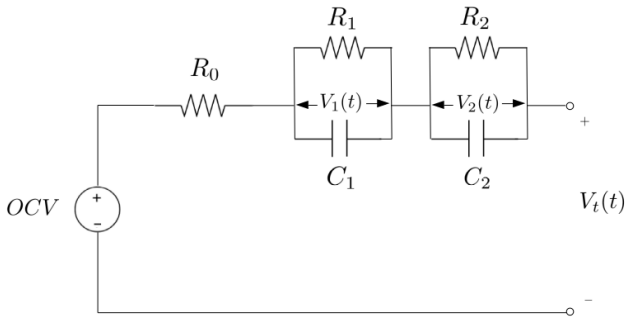

Figure 1. The battery model.

The following differential equations govern the behavior of the second-order circuit model in Figure 1,

$$\frac{dz(t)}{dt} = \frac{I(t)}{C_b}, \tag{1}$$

$$\frac{dV_1(t)}{dt} = -\frac{V_1(t)}{R_1C_1} + \frac{I(t)}{C_1}, \tag{2}$$

$$\frac{dV_2(t)}{dt} = -\frac{V_2(t)}{R_2C_2} + \frac{I(t)}{C_2}, \tag{3}$$

$$V_t(t) = OCV + V_1(t) + V_2(t) + R_0I(t), \tag{4}$$

where $I(t)$ indicates the current passing through the cell and $I(t) > 0$ when the cell is being charged, $z(t)$ denotes cell's state of charge (SOC), C_b is the capacity of the cell. Equation (1) expresses the relationship between current and SOC, where positive current increases SOC. Furthermore, (2) and (3) calculate the value of the polarization voltages $V_1(t)$ and $V_2(t)$. Lastly, (4) yields the cell's terminal voltage using the cell's open circuit voltage (OCV). Note that R_0, R_1, R_2, C_1, and C_2 change by the cell's core temperature as well as SOC. On the other hand, OCV is a function of the cell's SOC [13].

Thermal model

A two-state thermal model is employed to describe the thermal behavior of each cell. Further, the cell-to-cell heat exchange is also integrated into the model as follows,

$$\frac{dT_c(t)}{dt} = \frac{T_s(t) - T_c(t)}{R_cC_c} + \frac{\left|I(t)\left(OCV - V_t(t)\right)\right|}{C_c}, \tag{5}$$

$$\frac{dT_s(t)}{dt} = \frac{T_f(t) - T_s(t)}{R_uC_s} - \frac{T_s(t) - T_c(t)}{R_cC_s} - \frac{1}{R_mC_s}\sum_{i=1}^{N}(T_s(t) - T_i(t)), \tag{6}$$

where $T_f(t)$, $T_c(t)$, $T_s(t)$, and $T_i(t)$ are ambient, cell's core, cell's surface temperatures, and the surface temperature of neighbor cells, respectively. Equations (5) and (6) show the cell's core and surface temperature changes, respectively. The rate of core temperature change in (5) is controlled by conduction resistance R_c and core heat capacity C_c. On the other hand, surface temperature alteration in (6) is influenced by convection resistance R_u, cell-to-cell heat transfer resistance R_m, as well as surface heat capacity C_s. Additionally, each cell in the pack has N neighbor cells that affect the surface temperature of the cell in question.

Aging model

The aging model characterizes the cell capacity loss through the following lumped equation [14],

$$\Delta Q(t) = Be^{-\frac{E^a}{RT}}A_h(t)^\gamma \tag{7}$$

$$E^a = -31700 + 370.3 \times \text{C-rate} \tag{8}$$

$$A_h(t) = \int_{\tau=0}^{t} |I(\tau)|d\tau \tag{9}$$

$$B = \begin{cases} 31630, & C/2 \\ 21681, & 2C \\ 12934, & 6C \\ 15512, & 10C \end{cases} \tag{10}$$

where $\Delta Q(t)$ represents the percentage of capacity loss, E^a is the activation energy, T is the cell's core temperature, $R = 8.314$ [J/K/mol] is the universal gas constant, $A_h(t)$ shows cell ampere-hour throughput, and $\gamma = 0.55$. Moreover, (10) presents the values of B as a function of the C-rate. If the C-rate value is something other than $C/2$, $2C$, $6C$, or $10C$, parameter B is interpolated. In addition, we limit the output action of the RL agent to 0.5C and 10C such that extrapolation is not required.

Methodology

This section outlines the states, action, and reward of the proposed deep RL algorithm.

Markov decision process

Markov decision process (MDP) is a technique for modeling sequential decision-making problems where control actions can partly control the results. An MDP sequence is as follows. The system in question or the environment is at state s. Then the agent takes action a that takes the system from state s to state s' with the probability $p(s', r|s, a)$ and generates reward $r(s, a, s')$. When MDP is used for an optimization problem, the goal is to maximize the cumulative rewards. There are several approaches to solving an MDP, such as dynamic programming and RL.

Environment

In our problem, the battery pack is considered as the environment. The pack consists of N_s series connected batches of N_p parallel cells. In this study, $N_s = 6$ and $N_p = 74$, which is the configuration used in Tesla battery modules. Further, the considered lithium-ion cells are A123 ANR26650M1. The RL agent interacts with the environment by taking actions. The actions change the environment states according to (1)-(10), and the environment outputs a reward which is used to train the deep RL.

Observation space

We consider two states for the RL in this study: the surface temperature of the hottest cell in the pack and the SOC. We assume that the pack is equipped with a BMS that keeps the cell's SOCs balanced. As for the highest surface temperature, we assume that the temperature is within [5℃, 45℃] as these are the temperature limits for which the data of the parameter values in (1)-(4) is obtained [13]. In addition, the hottest surface temperature is at the center of the pack due to the impact of neighbor cells. Lastly, the SOC is clearly limited to [0,1].

Action space

The agent takes action to maximize the cumulative rewards. Hence, in our problem, the charging current is the action that should be determined to change the system's state, i.e., SOC. However, for simplicity, we train the RL to output the current C-rate rather than the current. As mentioned previously, the C-rate is limited to [0.5C, 10C].

Reward

The reward signal $r(s, a, s')$ indicates the desirability of transitioning from state s to s' by taking action a. In other words, the reward expresses the objective function of the optimization problem, which is defined as follows,

$$\begin{aligned} R(t) = -\omega_1 |z(t) - z^{max}| \\ - \omega_2 \max\{T_c^{max}(t) - T_c^{min}, 0\} \\ - \omega_3 \Delta Q^{max}(t), \end{aligned} \tag{11}$$

where $z^{max} = 0.8$ is the target SOC level at which the charging process stops, $T_c^{max}(t)$ denotes the core temperature of the hottest cell in the pack, T_c^{min} indicates the minimum core temperature above which the RL is penalized, and $\omega_i, i \in \{1,2,3\}$ is the weighting factor. The reward function is defined in such a way to be non-positive, i.e., $R(t) \leq 0$.

Results and analysis

In this section, the performance of the proposed model-free RL is evaluated and benchmarked against two different CC-CV charging profiles with different constant current stages of 4C and 5C, respectively. Moreover, every charging cycle charges the battery from 20% to 80% of SOC. The battery comprises 444 cells, including six series strings of 74 parallel cells. In addition, two tests are performed. In the first one, the battery pack is fast charged for one cycle, and the process is repeated three times, once with the proposed RL strategy, once with the CC-CV with 4C constant current stage, and once with the CC-CV with 6C constant current stage. In this test, we assume that a fresh battery is being charged for the very first time. Thus, at the beginning of the charging session, the battery's ampere-hour throughput is zero or all three cases. In the second test, the fresh battery pack is charged and discharged for 1000 cycles using the proposed approach and the two abovementioned CC-CV algorithms. Consequently, the first test is used to mainly investigate the temperature rise and speed of charging in each case. On the other hand, the second test reveals the cell aging after 1000 charge/discharge cycles.

Figure 2 and Figure 3 illustrate the results of charging the battery pack with the proposed methodology. Initially, the battery is soaked at the ambient temperature. As seen, the charging session takes 425 seconds to end. Since the number of cells is enormous, the legend is in the form of a color bar. The charging current is monotonically decreasing from 5.32C to 3.96C. Furthermore, the core temperature of the cells increases from the ambient temperature to around 36℃ with maximum and minimum core temperatures of 36.77℃ and 35.10℃, respectively. Similarly, the surface temperature of the cells starts from the ambient temperature to around 33℃ with maximum and minimum cell surface temperatures of 34.61℃ and 32.68℃, respectively. The surface temperature gradient is higher than that of core temperature, seeing that the surface of the cells is directly in contact with the ambient. In addition, the variations in the cells' charging current are visible due to temperature differences. Further, the cells' capacities have dropped by 0.15%. It merits mentioning that the capacity fade is exponential. At the early charging cycles, the capacity drop is quicker. As the cell ages, the reduction of capacity per cycle slows down. After the first cycle, the maximum and minimum cell capacity drop are 0.1531% and 0.1433%, respectively. Finally, Figure 3 shows the heatmap of the cells at the end of the charging process, where bright and dark colors indicate higher and lower temperatures, respectively.

Figure 4 depicts the results of charging the battery pack by CC-CV with a constant current stage of 4C. Compared to the results of the proposed approach in Figure 4, the core and surface temperatures are slightly lower, but 525 seconds is required to finish the charging session and reach 80% of SOC. The maximum and minimum cell core temperatures are 36.13℃ and 34.50℃, respectively, while the maximum and minimum cell surface temperatures are 34.32℃ and 32.23℃, respectively. On average, the core and surface temperatures of the cells are 0.495℃ less than those of PPO results.

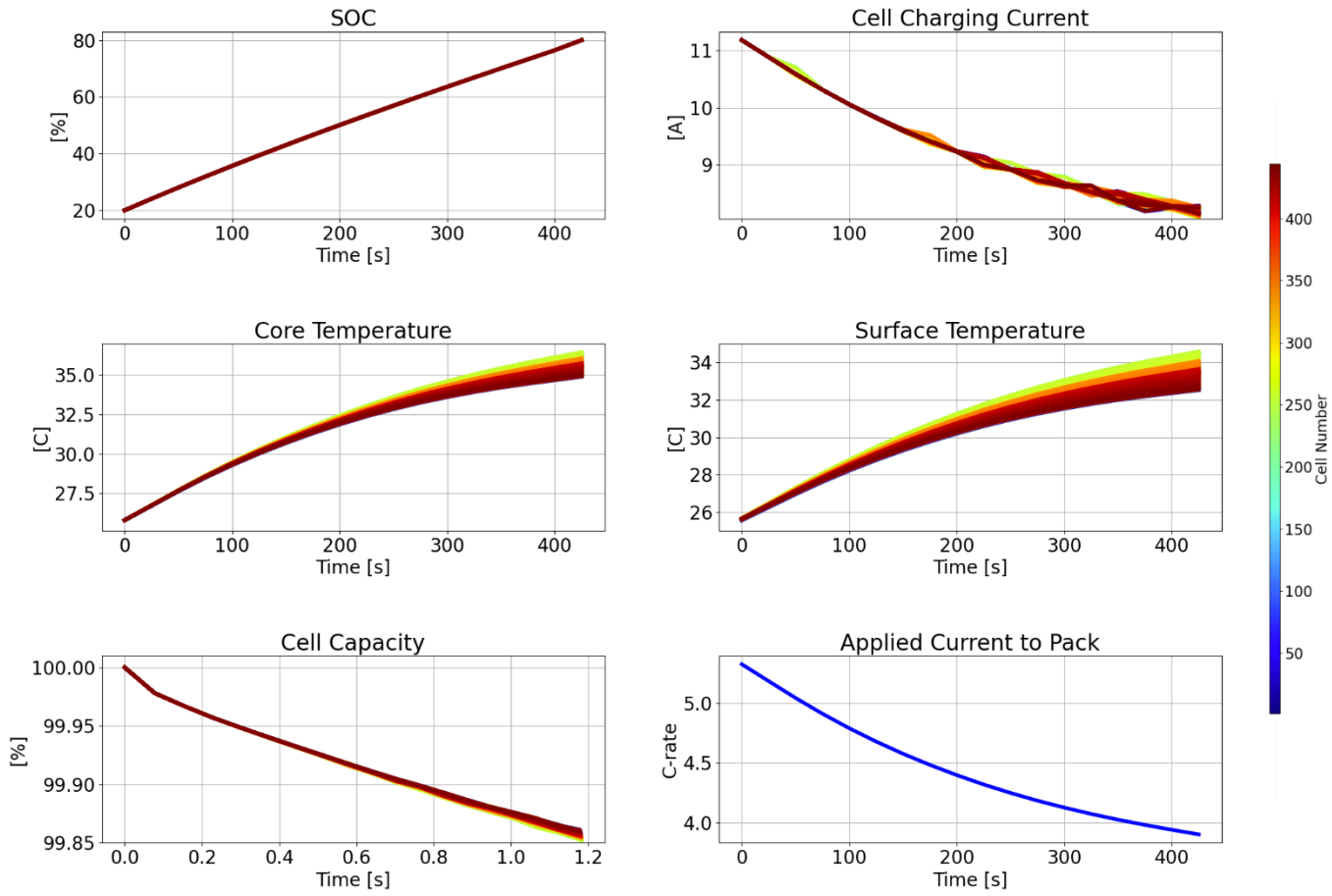

Figure 2. The results of charging the battery pack with the proposed approach based using PPO.

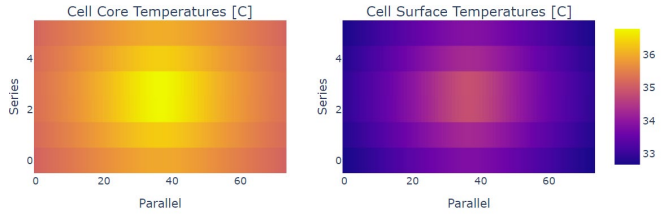

Figure 3. The heatmap of the cells core and surface temperatures. The bright color indicates hot cells, as opposed to the dark color, which shows cells with lower temperatures.

Figure 5 displays the results of charging the pack through a CC-CV with a 5C constant current stage. Identical to the proposed approach, Figure 5 shows the charging session ends after 425 seconds. However, the temperatures are elevated, and the degradation deteriorates. The cell core maximum and minimum temperatures rise to 39.01℃ and 37.30℃, respectively, and the surface maximum and minimum temperatures at the end of the charging session are 36.58℃ and 34.35℃, respectively. Consequently, on average, the cells are 2.02℃ hotter compared to the PPO case.

Lastly, Figure 6 exhibits average cell capacity change after 1000 fast charging cycles for PPO and CC-CV with 5C constant current stage. The result of PPO is compared with that of CC-CV with 5C as the constant current stage seeing that both approaches charge the battery in 425 seconds. In the case of PPO, the average cell capacity drops from 2.1 to 1.9614 [Ah], while the capacity drops to 1.9530 [Ah] if the battery is charged with the said CC-CV approach. In other words, in the case of PPO, the capacity drop was 6.6%, as opposed to the 7% drop in the CC-CV case. It is worth noting that a 0.4% improvement in capacity reduction is roughly equivalent to 60 charging sessions assuming the vehicle is charged once a week on average. Therefore, the lifetime of the battery pack is extended by about one year, as there are 52 weeks in a year. Hence, the proposed deep RL-based approach can charge the battery as fast as a CC-CV with a 5C constant current stage and mitigate battery degradation.

Conclusions

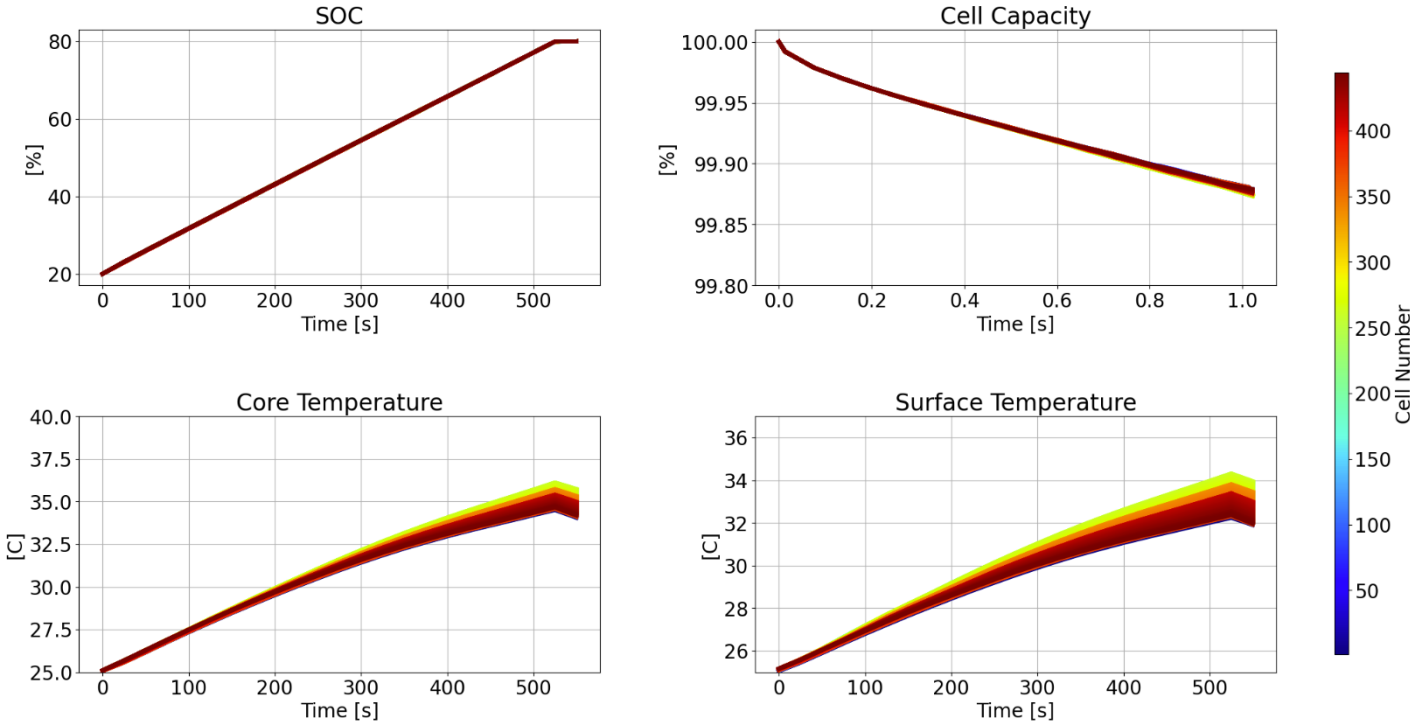

Figure 4. The results of charging the battery pack through CC-CV approach with 4C constant current stage.

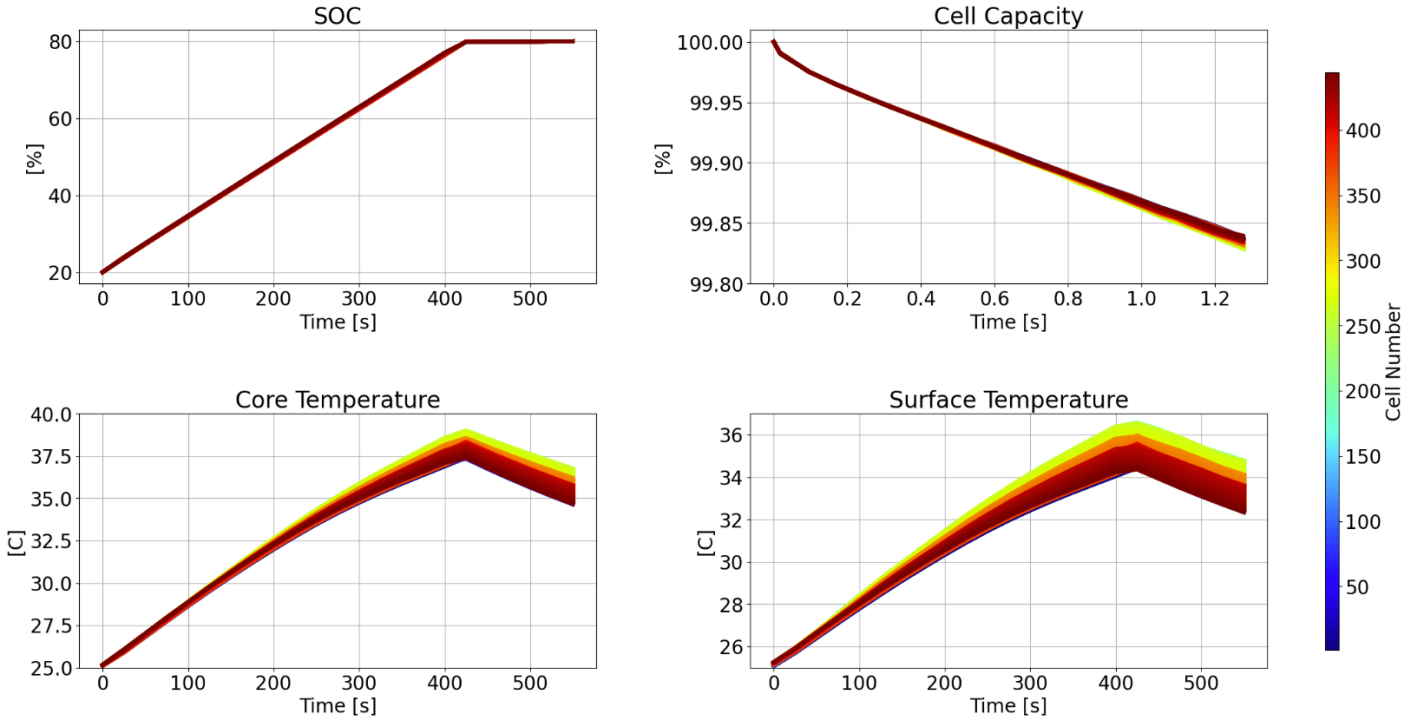

Figure 5. The results of charging the battery pack through CC-CV approach with 5C constant current stage.

This paper proposes a novel approach to fast charge a large battery pack through deep RL. First, the battery pack is modeled utilizing a second-order equivalent circuit model, a two-state thermal model which accounts for the cell-to-cell heat transfer, and a lumped aging model that describes the capacity fade of the battery. The model is used to characterize each battery pack cell, while the pack comprises 444 A123 ANR26650M1 lithium-ion cells. In addition, the cells are arranged in six series strings of 74 parallel cells. The underlying deep

Contact Information: {mabbasi, jiangfz}@clemson.edu

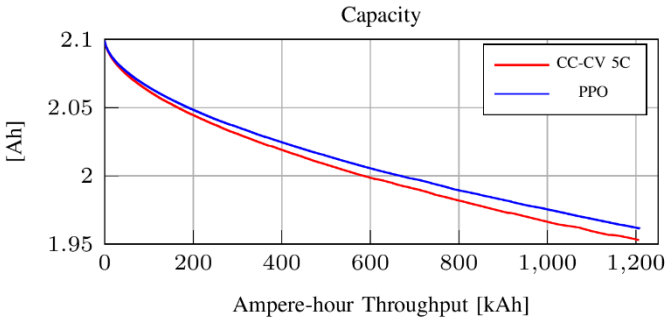

Figure 6. Capacity drop after 1000 fast charging cycles.

The PPO algorithm trains RL. Ultimately, the well-trained agent can charge the battery as fast as a CC-CV with a 5C constant current stage while keeping the average cell temperature 2.02℃ cooler. Furthermore, after 1000 fast charging cycles, the average capacity fade of the cells is 0.4% less in the RL-based charging case than the CC-CV with a 5C constant current stage. Moreover, the proposed RL methodology is also compared with a CC-CV with a 4C constant current stage. The results demonstrate that PPO charges the battery pack 100 seconds faster while the average battery temperature increases by 0.495℃.

Acknowledgments

This material is based upon work supported in part by the National Science Foundation under Grant No. CMMI-2312196.

References

1. Bilal, M., Ahmad, F., & Rizwan, M. (2023). Techno-economic assessment of grid and renewable powered electric vehicle charging stations in India using a modified metaheuristic technique. Energy Conversion and Management, 284, 116995.

2. Aretxabaleta, I., De Alegria, I. M., Andreu, J., Kortabarria, I., & Robles, E. (2021). High-voltage stations for electric vehicle fast-charging: trends, standards, charging modes and comparison of unity power-factor rectifiers. IEEE Access, 9, 102177-102194.

3. Abbasi, M. H., Zhang, J., & Krovi, V. (2022, November). A Lyapunov Optimization Approach to the Quality of Service for Electric Vehicle Fast Charging Stations. In 2022 IEEE Vehicle Power and Propulsion Conference (VPPC) (pp. 1-6). IEEE.

4. Tan, X., Lyu, P., Fan, Y., Rao, J., & Ouyang, K. (2021). Numerical investigation of the direct liquid cooling of a fast-charging lithium-ion battery pack in hydrofluoroether. Applied Thermal Engineering, 196, 117279.

5. Abbasi, M. H., & Zhang, J. (2021, December). Joint optimization of electric vehicle fast charging and dc fast charging station. In 2021 International Conference on Electrical, Computer and Energy Technologies (ICECET) (pp. 1-6). IEEE.

6. Park, S., Pozzi, A., Whitmeyer, M., Perez, H., Kandel, A., Kim, G., ... & Moura, S. (2022). A deep reinforcement learning framework for fast charging of li-ion batteries. IEEE Transactions on Transportation Electrification, 8(2), 2770-2784.

7. Pozzi, A., & Raimondo, D. M. (2022). Stochastic model predictive control for optimal charging of electric vehicles battery packs. Journal of Energy Storage, 55, 105332.

8. Chen, S., Bao, N., Garg, A., Peng, X., & Gao, L. (2021). A fast charging–cooling coupled scheduling method for a liquid cooling-based thermal management system for lithium-ion batteries. Engineering, 7(8), 1165-1176.

9. Thakur, A. K., Sathyamurthy, R., Velraj, R., Saidur, R., Pandey, A. K., Ma, Z., ... & Ali, H. M. (2023). A state-of-the art review on advancing battery thermal management systems for fast-charging. Applied Thermal Engineering, 226, 120303.

10. Wei, Z., Quan, Z., Wu, J., Li, Y., Pou, J., & Zhong, H. (2021). Deep deterministic policy gradient-DRL enabled multiphysics-constrained fast charging of lithium-ion battery. IEEE Transactions on Industrial Electronics, 69(3), 2588-2598.

11. Hao, Y., Lu, Q., Wang, X., & Jiang, B. (2023). Adaptive Model-Based Reinforcement Learning for Fast Charging Optimization of Lithium-Ion Batteries. IEEE Transactions on Industrial Informatics.

12. Yang, Y., He, J., Chen, C., & Wei, J. (2023). Balancing Awareness Fast Charging Control for Lithium-Ion Battery Pack Using Deep Reinforcement Learning. IEEE Transactions on Industrial Electronics.

13. Lin, X., Perez, H. E., Mohan, S., Siegel, J. B., Stefanopoulou, A. G., Ding, Y., & Castanier, M. P. (2014). A lumped-parameter electro-thermal model for cylindrical batteries. Journal of Power Sources, 257, 1-11.

14. Wang, J., Liu, P., Hicks-Garner, J., Sherman, E., Soukiazian, S., Verbrugge, M., ... & Finamore, P. (2011). Cycle-life model for graphite-LiFePO4 cells. Journal of power sources, 196(8), 3942-3948.