# Intelligent Control for Type I Partial Power Converters in EV Charging Systems
## Twin-Delayed Deep Deterministic Policy Gradient Approach

Pesantez, Daniel; Menendez, Oswaldo; Renaudineau, H.; Kouro, S.; Rivera, S.; Rodriguez, Jose

**DOI**
10.1109/ICA-ACCA62622.2024.10766834

**Publication date**
2024

**Document Version**
Final published version

**Published in**
2024 IEEE International Conference on Automation/26th Congress of the Chilean Association of Automatic Control, ICA-ACCA 2024

2024 IEEE International Conference on Automation/XXVI Congress of the Chilean Association of Automatic Control (ICA-ACCA) | 979-8-3503-7811-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICA-ACCA62622.2024.10766834

# Intelligent Control for Type I Partial Power Converters in EV Charging Systems: Twin-Delayed Deep Deterministic Policy Gradient Approach

Daniel Pesantez
*Centro de Transición Energética, Facultad de Ingeniería, Arquitectura y Diseño, Universidad San Sebastián.*
Santiago, Chile
alvaro.pesantez@sansano.usm.cl

Oswaldo Menendez
*Departamento de Ingeniería de Sistemas, y Computación, Universidad Católica del Norte.*
Antofagasta, Chile
oswaldo.menendez@ucn.cl

H.Renaudineau
*Department of Electronic Engineering Univ. Tecnica Federico Santa Maria*
Valparaiso, Chile
hugues.renaudineau@usm.cl

S. Kouro
*Department of Electronic Engineering Univ. Tecnica Federico Santa Maria*
Valparaiso, Chile
samir.kouro@ieee.org

S. Rivera
*Department of Electrical Sustainable Energy Delft University of Technology*
Delft, The Netherlands
s.rivera.i@ieee.org

Jose Rodriguez
*Centro de Transición Energética, Facultad de Ingeniería, Arquitectura y Diseño, Universidad San Sebastián.*
Santiago, Chile
jose.rodriguezp@uss.cl

***Abstract*—In recent years, the electric vehicle (EV) industry has experienced significant advancements, simultaneously driving substantial progress in battery technology. The evolution of battery systems necessitates enhancements in charging infrastructure to attain elevated power levels during the charging process, thereby minimizing charging time. Various algorithms have been developed for driving battery charging; however, these algorithms necessitate the creation of diverse controllers to generate precise trigger signals for the semiconductors within the various power converters utilized in charging stations. This work presents the design of an innovative model-free control system for Type I impedance network Partial Power Converter (PPC) in which a Deep Reinforcement Learning (DRL) agent generates control signals during the different charging stages. Particularly, a Twin-Delayed Deep Deterministic Policy Gradient (TD3) algorithm is used to substitute the inner control loop of traditional control systems. To this end, different agents were designed, trained, and tested inside a built simulation environment. It is worth noting that TD3-based control allows for the optimal functionality of a type I impedance network PPC within the context of EV battery charging applications, according to the specified CC-CV charging algorithm. Empirical results revealed that the battery system reached an 80% state of charge in under 8 minutes starting from an initial 20%.**

***Index Terms*—neural network, deep reinforcement learning, electric vehicles, partial power converter, artificial intelligence**

## I. Introduction

One of the main worldwide challenges is related to reducing the levels of greenhouse gas emissions, particularly Carbon Dioxide ($CO_2$). The transportation industry accounts for 16% of $CO_2$ emissions due to the burning of fossil fuels in Internal Combustion Engines (ICE) [1], [2]. In the last years, the Electric Vehicle (EV) sector has emerged as a highly promising field, aiming to substitute ICE with fully electric-powered engines [3]. However, a significant challenge in this industry evolution has been battery capacity and operating range [4]. Consequently, advancements in battery technology over the past ten years have led to increased battery capacities and decreased costs [5].

Alongside advances in battery technology, the need for more efficient chargers has become essential, leading to extensive research on different AC-DC and DC-DC converter topologies [6]–[9]. Fig. 1 illustrates the general configuration of an off-board charging station, where all components involved in the charging process, except the battery, are located outside the EV. The charging process is divided into two stages: an initial AC-DC rectifier stage, followed by a DC-DC conversion stage. Galvanic isolation can be implemented in two ways. The first architecture, depicted in Fig. 1 a), involves a low-frequency tranformer at the AC grid side. The second architecture, shown in Fig. 1 b), conisders galvanic isolation in the DC-DC stage using a high-frequency transformer. Regarding the DC-DC conversion stage, numerous converters have been designed featuring different topologies [6], [8], [10], among which, Partial Power Converters (PPC) have become a viable option for the DC-DC conversion stage in charging systems [11], [12]. By handling only a portion of the total power, they reduce converter-associated losses, thus improving the overall efficiency of the system.

To perform the charging process for EV batteries, numerous algorithms have been developed. These aim to minimize the charging time and extend the life cycle. To achieve this, some of these algorithms take into account factors such as temperature, state of health, current injection rate, and various physical and chemical phenomena that impact the battery over

979-8-3503-7811-5/24/$31.00 ©2024 IEEE

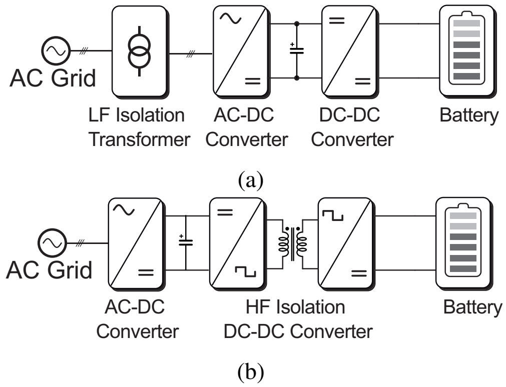


Fig. 1. General structure of an EV battery charger with: a) Low frequency isolation, b) High frequency isolation.

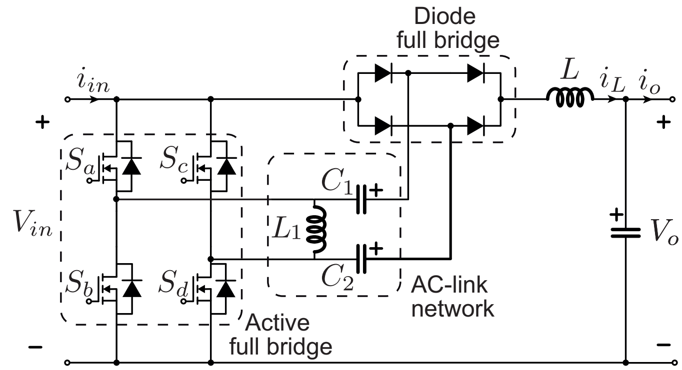


Fig. 2. Type I PPC. [11]

time and with repeated charge-discharge cycles [13]–[15]. For example, the Constant Current and Constant Voltage (CC-CV) charging algorithm [16] is the most commonly used in EV battery charging applications. This algorithm consists of two operating stages. In the initial CC stage, the battery is charged with a constant current, the maximum value of which is primarily set by the battery manufacturer and the charger capacity; the second CV stage, avoids overcharging the battery by limiting the current flowing into it, enabling the battery to achieve a full 100% State of Charge (SoC) in a regulated way. Another algorithm used for battery charging is to consider the pulsed profile, where the battery receives periodic pulses of current with varying amplitude and frequency [17]. A third charging algorithm described in the literature involves the use of multiple stages of constant current. This method aims to minimize the charging time by identifying the optimal current value to be supplied to the battery based on its SoC at each stage, thereby avoiding excessive temperature rise throughout the charging process [18], [19].

In the CC-CV charging algorithm, a cascade control system utilizing Proportional-Integral (PI) controllers is commonly used. In [20] a converter is presented that uses a non-linear input-output feedback linearization controller.

On the other hand, data-driven control systems are advanced control systems that determine the dynamics of power converter through simulated and experimental data [21]. Recently, Deep Reinforcement Learning (DRL) agents have gained more attention due to their high adaptability and robustness to disturbances and mismatched parameters, enabling the development of model-free adaptive control systems [22], [23]. Different DRL agents have been proposed as control strategies for different power converters. Two application fields lead the current approaches: **(i)** DC-DC [24]–[27] and **(ii)** DC-AC [28]–[33]. Continuous control actions in DC-DC power converters enable the use of actor-critic DRL agents, such as the Proximal Policy Gradient (PPG) [24], Deep Deterministic Policy Gradient (DDPG) [25] and Twin-Delayed Deep Deterministic Policy Gradient (TD3) [26]. On the other hand, a value-based DRL algorithm, called Deep Q-network (DQN) is proposed in [27]. Due to the structure of DQN, the proposed approach reduces the required neural networks.

Here, an innovative artificial intelligent control system based on a TD3 algorithm is introduced to control the charging process of EV battery using a type I impedance network PPC which as been proposed in [11]. The proposed approach replaces the internal current controller within a traditional cascade configuration, thus facilitating the regulation of the EV battery charging process based on the CC-CV charging algorithm. The main idea is to replace the traditional inner current control loop with an action-critic DRL agent, while maintaining the PI control as the external loop. The proposed control system is trained using a simulated environment to avoid potential damage or risks to the power converter structure. In addition, the behavior of the TD3-based control is evaluated under variation in control constants of the external control loop and the load system parameters. The remainder of this work is organized as follows. Section II delineates the basic principles associated with the type I impedance network PPC used for the charging process. Section III reviews the mathematical formulation of the TD3-based control system and presents the design of the TD3-agent. Section IV shows the main simulation findings derived from the charging process and the training of the TD3-agent. Finally, Section V details the conclusions of this work.

## II. Type I impedance network PPC

The behavior of a data-driven control system based on the TD3 algorithm is assessed in a Type I impedance network PPC [11]. Fig. 2 illustrates the power converter topology. Note that the power converter consists of an active full bridge connected to a diode full bridge through an impedance network, which includes the inductor $L1$ and the capacitors $C_1$ and $C_2$. This topology generates a series DC voltage between the input voltage $V_{in}$ of the DC-DC converter and the battery.

Moreover, the switching signals $S_a$, $S_b$, $S_c$, and $S_d$ are generated using the phase shift modulation strategy (PSM). This approach enables a phase shift in the voltage produced by each leg of the active full bridge. According to [11], the shift value or displacement factor $\alpha$ ranges from 0 to 0.5. The voltage across the output inductor $L$ is described as follows:

$$L\frac{d\bar{i}_L}{dt} = \alpha(\bar{V}_o - V_{in}) + (\bar{V}_o - 2V_{in})(1 - \alpha), \quad (1)$$

where $\bar{x}$ denotes the average value of any signal $x$ during a switching period $T$. $i_L$ denotes the current flowing through the inductor $L$.

Assuming the system is in steady-state, i.e., $\frac{d\bar{i}_L}{dt} = 0$, the input-to-output transfer function of the system can be derived as follows,

$$\frac{\bar{V}_o}{V_{in}} = 2 - \alpha, \quad (2)$$

as shown in 2, the output voltage of the power converter is directly dependent on $\alpha$, allowing an voltage increase between 1.5 and 2 (for more information, see [11]).

## III. Intelligent control system design

A TD3 agent is an off-policy, model-free DRL technique that consists of policy, reward function, action-value function, and environment [34]. This work presents a data-driven control system based on TD3 algorithm, following the guidelines described in [35]. Fig. 3 illustrates the methodology used in this work. The data-driven control system will be referred to as the TD3-based control for the remainder of this work. The TD3-based control can be summarized as follows.

- The action and observation spaces are built based on the power converter control philosophy. The observations, $x_t$, consist of output load inductor current $i_L$, the current error $e_{i_L}$, the integral of the error, and the voltage error in the battery $e_{V_o}$. The action space comprises the displacement factor $\alpha$. To guarantee the stability of $V_o$, the TD3 agent is trained using a PI control that computes the current reference $i_L^*$.
- The TD3 algorithm consists of two critic networks and one actor network, as shown in Fig. 3. To avoid vanishing and exploding gradient problems related to large neural networks, this work employs a single hidden layer actor network with 100 neurons. Since $\alpha$ has a continuous behavior, a hyperbolic tangent activation function is selected. To reduce the exploration process of the TD3 agent, a scaling stage is added that modifies the range of $\alpha$ form $[-1, 1]$ to $[0, 0.5]$. Finally, a PSM is implemented using the guidelines described in [11]. On the other hand, critic networks consist of two hidden layers, each with 100 neurons. They are also equipped with rectified linear unit (RELU) activation functions. All neural networks are designed using Deep Network Designer toolbox.
- The agent must first be trained in a controlled and safe simulation environment to ensure the convergence of the TD3-based control system and to reduce the potential damages in the power converter or battery system. Thus, a simulation environment is created using MATLAB 2023/Simulink, which emulates the behavior of the power converter, battery system, and measurement system. During the training, the environment modifies the reference voltage in the battery $V_o^*$ for four different values between 610 V and 790 V. The TD3 agent is created using the Reinforcement Learning Toolbox and is trained and tested using a computer (Intel(R) Core(TM) i7-10700F processor running at 2.90GHz, with eight cores). Furthermore, the computer has an NVIDIA GeForce RTX 3070 GPU and 32 GB of RAM.
- The fundamental concept behind TD3 is associated with the reward hypothesis, which states that all objectives and purposes can be framed as maximizing the expected cumulative reward. To this end, we propose a reward function that minimizes $e_{i_L}$, $e_{V_o}$, and the control effort using the previous time step action $\alpha_{k-1}$. The candidate reward function for each time step is defined as follows:

$$r_k = -\left(Q_1 \left|i_L^* - i_L\right| + Q_2 \left|V_o^* - V_o\right| + Q_3 \left|\alpha_{k-1}\right|\right), \quad (3)$$

  where $Q_1 = Q_2 = 20$, and $Q_3 = 5$ are constants, empirically determined.
- After training the agent, the control system is adapted to follow the CC-CV charging algorithm described in [11]. Due to the independence between the TD3 agent and the constants of the external PI controller, it is possible to tune control parameters without re-training the agent. The battery charging process consists of two main stages. First, in the CC stage, the converter is driven by a single current-loop control system, i.e., the reference for the internal current controller is set to a constant value indicated by the BMS. During this stage, the $e_{V_o}$ is set to zero. Once the battery voltage reaches a SoC of 80%, a bypass-control strategy autonomously commutes to a cascade control system in the CV stage. In other words, the $e_{V_o}$ is connected to an external PI voltage controller. The external PI controller generates the current reference to ensure the battery system reaches a full SoC in a controlled manner.

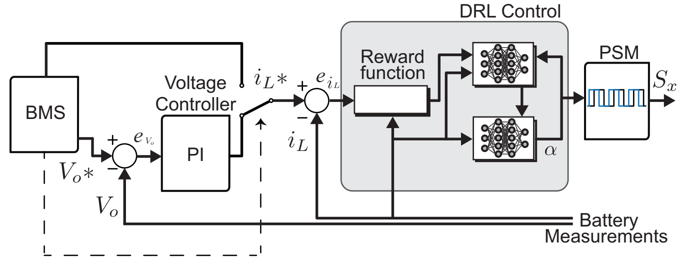

Fig. 3. Proposed control scheme for the battery charging application.

## IV. Simulation Results

This section analyses the performance of TD3-based control for driving a Type I impedance network PPC.

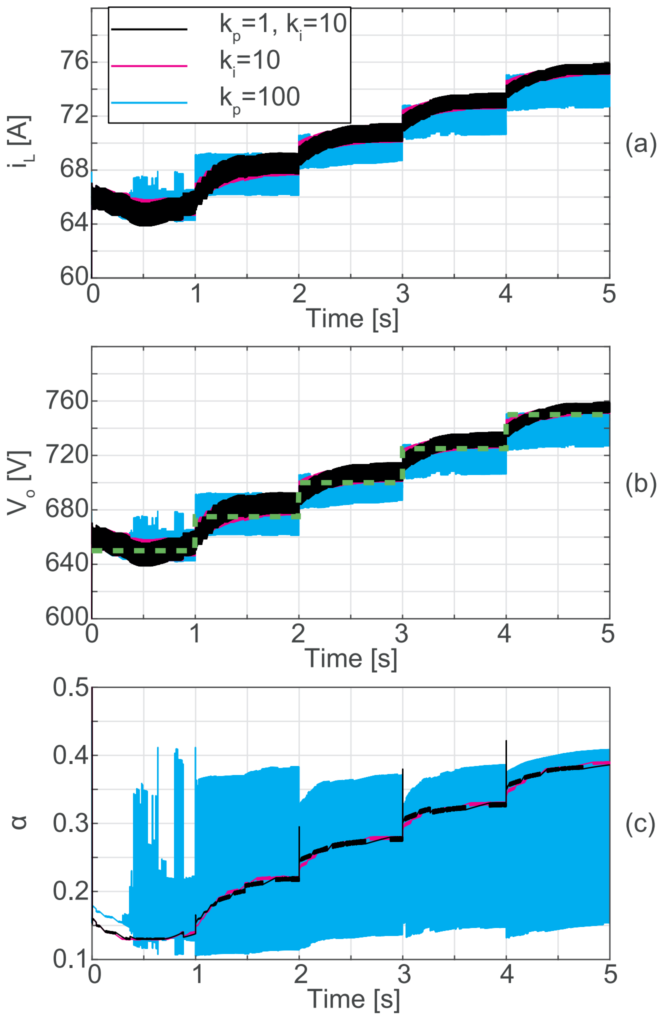


Fig. 4. Performance of trained TD3 agent under variation in $k_p$ and $k_i$: a) Load inductor current. b) Voltage in the DC-link. c) Displacement factor.

TABLE I
SIMULATION PARAMETERS

| **Parameters** | **Value** |
|---|---|
| Input Voltage | 400 V |
| CC mode | 100 A |
| Output inductor | 1 mH |
| L1 inductor | 1 mH |
| $C_1$ and $C_2$ capacitors | 200 uF |
| Switching frequency | 20 kHz |

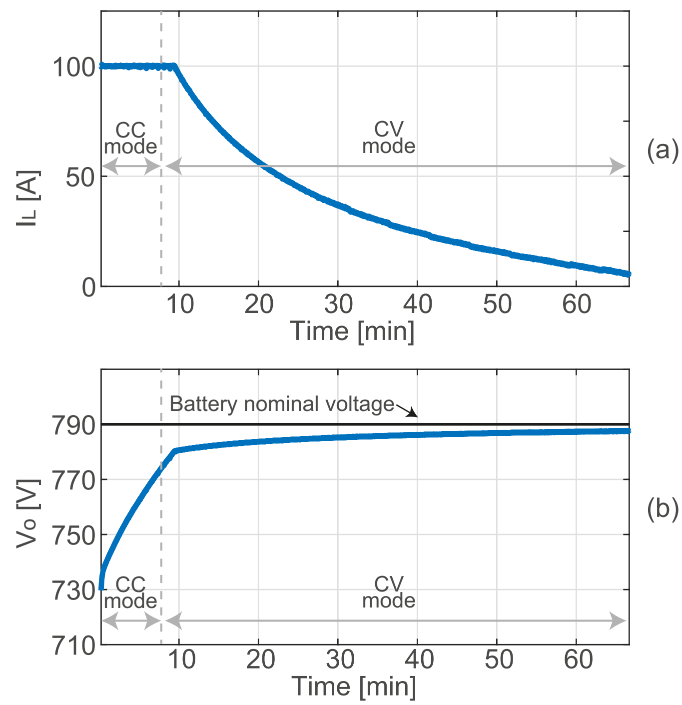


Fig. 5. Battery charger simulation results: a) Output converter current. b) Battery voltage.

### *A. Training of TD3-agent*

The TD3 agent designed in Section III is trained using the Reinforcement Learning Toolbox. To ensure proper convergence, the TD3 agent is trained using a sampling time of 20 $\mu$s over 20 episodes, each lasting 2 s, with the agent frequency set at 50 kHz. For the sake of this analysis, we also use a 10 $\Omega$ as a load during the training stage. According to [35], the performance of the TD3 agent does not depend on proportional and integral gains (kp and ki). Here, kp and ki are set to reasonable values, one and ten, during the training stage. The empirical findings revealed that the agent converges after 325 minutes.

Fig. 4 shows the main results of the trained agent at different values of kp and ki for five different $V_o^*$. It is worth noting that once the agent is trained, it is robust enough to vary in the voltage control system. Although $i_L$ and $V_o$ can be controlled in three analyzed scenarios, the prediction $\alpha$ is poor to proportional control (see cyan continuous lines in Fig. 4c.) On the other hand, it is possible to see that the setting time of $i_L$ is less than 0.5 s (see Fig. 4a), which is sufficient for the charging battery process. Finally, the output current ripple decreases with the $\alpha$ increase.

### *B. Battery charger results*

To verify the proposed control method, the type I PPC is evaluated in an EV battery charging scenario, utilizing the converter parameters listed in Table I. To simulate the charging process, a battery model with a 60 kWh capacity and a nominal voltage of 790 V is utilized. Initially, the battery state of charge (SoC) is 20%, and it is charged to 100% using the CC-CV charging algorithm.

The simulation results of the complete charging procedure are illustrated in Fig. 5. The behavior of the charging algorithm is illustrated in Fig. 5 a), where the output current of the converter is depicted, showing that during the CC charging stage, a constant current of 100 A is supplied to the battery. In this phase, the intelligent controller receives a constant current reference until the battery reaches 80% of its SoC. At this

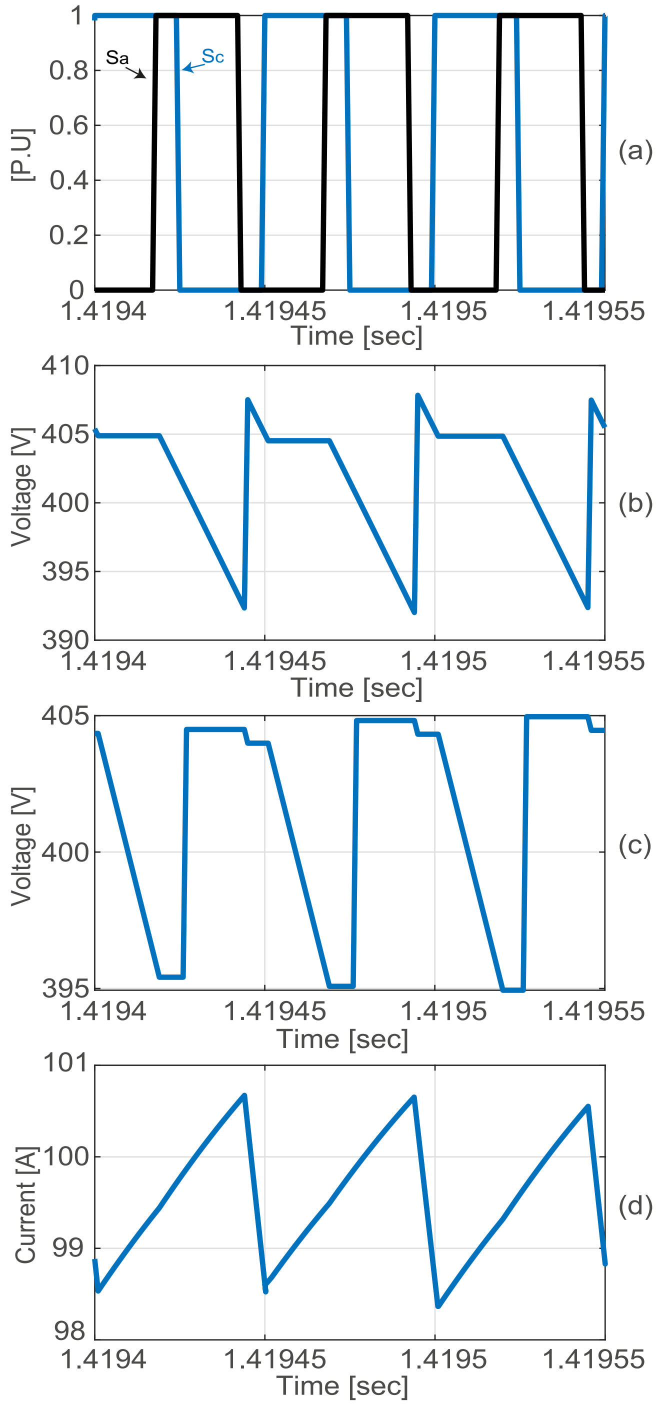


Fig. 6. Main waveforms of the converter observed during the battery charging simulation: a) $S_a$ and $S_c$. b) $C_1$ capacitor voltage. c) $C_2$ capacitor voltage. d) $i_L$.

point, the smart controller ceases to receive a constant current reference, transitioning to the CV charging stage. During this stage, an external PI voltage controller is tasked to provide the current reference. Initially, this signal matches the constant current reference from the previous phase until around minute 10, after which the current gradually decreases as the SoC approaches 100%. Fig. 5 b) illustrates the voltage behavior in the battery throughout the entire charging process. It indicates that the battery reaches 80% SoC when it reaches a voltage of 775 V, at which point it switches to the CV charging stage. During this stage, the voltage continues to increase until it reaches the nominal value of the battery. At approximately the 10-minute mark, as the charging current diminishes, the voltage ascends more gradually, ultimately achieving full charge in a controlled manner within an estimated duration of 60 min.

In steady state, during the CC charging mode, the voltage waveforms across the impedance network capacitors, as well as the current through the output inductor $i_L$, over three switching cycles, are shown in Fig. 6. As illustrated in Fig. 6 a), at this instant, the displacement factor $\alpha$ is equal to $0.14T$, where $T$ denotes a switching period. As depicted in Fig. 6 b) and c), the voltage across the capacitors in the impedance network in mean value is equal to the input voltage, remaining constant regardless of the output voltage of the converter. Fig. 6 d) shows the output current waveform of the converter, which shows a ripple of 2% at a current value of 100 A.

## V. Conclusions

This work presented an innovative TD3-based control for a type I impedance network PPC. Different TD3 algorithms were designed, trained, and tested in a simulated environment. The simulation demonstrates that the proposed controller could successfully charge a battery on the basis of the CC-CV charging protocol. The battery SoC increased from 20% to 80% in approximately 8 minutes. This test was carried out using a 60 kWh battery model and simulated a 79 kW charger. The simulation results also revealed that the adaptability of the strategy to the Type I impedance network PPC enables it to handle uncertainties, variations, and nonlinearities associated with the converter, resulting in improved performance and stability.

## Acknowledgment

This work was supported in part by AC3E (ANID/Basal/FB0008) and SERC Chile (ANID/Fondap/1523A0006). This work was supported in part by ANID-Fondecyt 1221741. This work was supported in part by Doctorado Nacional/2022/21221405.

## References

[1] U. E. P. Agency, "Fast facts on transportation greenhouse gas emissions," 2022, accessed: 2024-05-22. [Online]. Available: https://www.epa.gov/greenvehicles/fast-facts-transportation-greenhouse-gas-emissions

[2] J. Chen, S. Tao, S. Teng, Y. Chen, H. Zhang, and F.-Y. Wang, "Toward sustainable intelligent transportation systems in 2050: Fairness and eco-responsibility," *IEEE Transactions on Intelligent Vehicles*, vol. 8, no. 6, pp. 3537–3540, 2023.

[3] L. Dickerman and J. Harrison, "A new car, a new grid," *IEEE Power and Energy Magazine*, vol. 8, no. 2, pp. 55–61, 2010.

[4] L. Xia, C. Chen, H. Ren, and Z. Kang, "Analysis of range anxiety using nev monitoring big data," in *2022 Asia Conference on Algorithms, Computing and Machine Learning (CACML)*, 2022, pp. 145–149.

[5] BloombergNEF, "Lithium-ion battery pack prices hit record low of $139/kwh," 2023, accessed: 2024-05-23. [Online]. Available: https://about.bnef.com/blog/lithium-ion-battery-pack-prices-hit-record-low-of-139-kwh/

[6] S. Rivera, S. M. Goetz, S. Kouro, P. W. Lehn, M. Pathmanathan, P. Bauer, and R. A. Mastromauro, "Charging infrastructure and grid integration for electromobility," *Proceedings of the IEEE*, vol. 111, no. 4, pp. 371–396, 2023.

[7] F. Sedaghati and S. A. Azimi, "A non-isolated bidirectional dc-dc converter connected to the t-type three level converter for electric vehicle battery charger," in *2023 13th Smart Grid Conference (SGC)*, 2023, pp. 1–5.

[8] R. Singh, S. K. Gawre, and G. Dyanamina, "Review and analysis of dc-dc power converter performance for fast charging of evs," in *2021 IEEE 2nd International Conference On Electrical Power and Energy Systems (ICEPES)*, 2021, pp. 1–5.

[9] S. Chaurasiya and B. Singh, "An adaptive ev charger based reconfigurable bidirectional isolated dc-dc converter with wide voltage range of operation," in *2022 IEEE Transportation Electrification Conference & Expo (ITEC)*, 2022, pp. 1301–1306.

[10] S. Sarkar and A. Das, "An isolated single input-multiple output dc–dc modular multilevel converter for fast electric vehicle charging," *IEEE Journal of Emerging and Selected Topics in Industrial Electronics*, vol. 4, no. 1, pp. 178–187, 2023.

[11] D. Pesantez, H. Renaudineau, S. Rivera, A. Peralta, A. Marquez Alcaide, and S. Kouro, "Transformerless partial power converter topology for electric vehicle fast charge," *IET Power Electronics*, vol. n/a, no. n/a. [Online]. Available: https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/pel2.12613

[12] S. Rivera, D. Pesantez, S. Kouro, and P. W. Lehn, "Pseudo-partial-power converter without high frequency transformer for electric vehicle fast charging stations," in *2018 IEEE Energy Conversion Congress and Exposition (ECCE)*, 2018, pp. 1208–1213.

[13] X. Han, L. Lu, Y. Zheng, X. Feng, Z. Li, J. Li, and M. Ouyang, "A review on the key issues of the lithium ion battery degradation among the whole life cycle," *eTransportation*, vol. 1, p. 100005, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2590116819300050

[14] P. Keil and A. Jossen, "Charging protocols for lithium-ion batteries and their impact on cycle life—an experimental study with different 18650 high-power cells," *Journal of Energy Storage*, vol. 6, pp. 125–141, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352152X16300147

[15] X. Lin, K. Khosravinia, X. Hu, J. Li, and W. Lu, "Lithium plating mechanism, detection, and mitigation in lithium-ion batteries," *Progress in Energy and Combustion Science*, vol. 87, p. 100953, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360128521000514

[16] W. Shen, T. T. Vo, and A. Kapoor, "Charging algorithms of lithium-ion batteries: An overview," in *2012 7th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, 2012, pp. 1567–1572.

[17] K. S. Venkat, M. Satya Sai Chandra, and S. Mohapatro, "Pulse charging scheme for multiple battery charging in electric vehicle applications," in *2023 IEEE 3rd International Conference on Smart Technologies for Power, Energy and Control (STPEC)*, 2023, pp. 1–6.

[18] M. Usman Tahir, A. Sangwongwanich, D.-I. Stroe, and F. Blaabjerg, "Overview of multi-stage charging strategies for li-ion batteries," *Journal of Energy Chemistry*, vol. 84, pp. 228–241, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2095495623003091

[19] Y. Zhang, S. Xu, and T. Wu, "Multi-stage constant current charging strategy considering soc intervals and voltage thresholds," *Global Energy Interconnection*, vol. 5, no. 2, pp. 143–153, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2096511722000329

[20] T. Sojoudi, M. Sarhangzadeh, J. Olamaei, and J. F. Ardashir, "An extendable bidirectional high-gain dc–dc converter for electric vehicle applications equipped with iofl controller," *IEEE Transactions on Power Electronics*, vol. 38, no. 8, pp. 9767–9779, 2023.

[21] S. Zhao, F. Blaabjerg, and H. Wang, "An overview of artificial intelligence applications for power electronics," *IEEE Transactions on Power Electronics*, vol. 36, no. 4, pp. 4633–4658, 2021.

[22] X. Chen, G. Qu, Y. Tang, S. Low, and N. Li, "Reinforcement learning for selective key applications in power systems: Recent advances and future challenges," *IEEE Transactions on Smart Grid*, vol. 13, no. 4, pp. 2935–2958, 2022.

[23] S. Zhang, O. Wallscheid, and M. Porrmann, "Machine learning for the control and monitoring of electric machine drives: Advances and trends," *IEEE Open Journal of Industry Applications*, vol. 4, pp. 188–214, 2023.

[24] M. Hajihosseini, M. Andalibi, M. Gheisarnejad, H. Farsizadeh, and M.-H. Khooban, "Dc/dc power converter control-based deep machine learning techniques: Real-time implementation," *IEEE Transactions on Power Electronics*, vol. 35, no. 10, pp. 9971–9977, 2020.

[25] M. Gheisarnejad, H. Farsizadeh, and M. H. Khooban, "A novel nonlinear deep reinforcement learning controller for dc–dc power buck converters," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 8, pp. 6849–6858, 2021.

[26] Y. Tang, W. Hu, D. Cao, N. Hou, Z. Li, Y. W. Li, Z. Chen, and F. Blaabjerg, "Deep reinforcement learning aided variable-frequency triple-phase-shift control for dual-active-bridge converter," *IEEE Transactions on Industrial Electronics*, vol. 70, pp. 10 506–10 515, 2023.

[27] C. Cui, T. Yang, Y. Dai, C. Zhang, and Q. Xu, "Implementation of transferring reinforcement learning for dc-dc buck converter control via duty ratio mapping," *IEEE Transactions on Industrial Electronics*, vol. 70, pp. 6141–6150, 6 2023.

[28] S. Bhattacharjee, S. Halder, Y. Yan, A. Balamurali, L. V. Iyer, and N. C. Kar, "Real-time sil validation of a novel pmsm control based on deep deterministic policy gradient scheme for electrified vehicles," *IEEE Transactions on Power Electronics*, vol. 37, no. 8, pp. 9000–9011, 2022.

[29] O. E. Egbomwan, S. Liu, and H. Chaoui, "Twin delayed deep deterministic policy gradient (td3) based virtual inertia control for inverter-interfacing dgs in microgrids," *IEEE Systems Journal*, vol. 17, no. 2, pp. 2122–2132, 2023.

[30] B. She, F. Li, H. Cui, H. Shuai, O. Oboreh-Snapps, R. Bo, N. Praisuwanna, J. Wang, and L. M. Tolbert, "Inverter pq control with trajectory tracking capability for microgrids based on physics-informed reinforcement learning," *IEEE Transactions on Smart Grid*, 2023.

[31] Y. Wang, S. Fang, and D. Huang, "An improved model-free active disturbance rejection deadbeat predictive current control method of pmsm based on data-driven," *IEEE Transactions on Power Electronics*, vol. 38, no. 8, pp. 9606–9616, 2023.

[32] C. Xiang, X. Zhang, T. Qie, T. K. Chau, J. Ye, Y. Yu, H. H. C. Iu, and T. Fernando, "A novel deep deterministic policy gradient assisted learning based control algorithm for three-phase dc/ac inverter with an rl load," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, pp. 1–1, 2022.

[33] O. Menéndez, D. López-Caiza, L. Tarisciotti, F. Ruiz, F. Auat-Cheein, and J. Rodríguez, "Assessment of deep reinforcement learning algorithms for three-phase inverter control," in *2023 IEEE 8th Southern Power Electronics Conference and 17th Brazilian Power Electronics Conference (SPEC/COBEP)*, 2023, pp. 1–8.

[34] 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," 2019.

[35] O. Menendez, D. Lopez-Caiza, A. Prado, F. Flores-Bahamonde, and J. Rodríguez, "Intelligent control of an active front-end converter: Deep reinforcement learning approach," in *2023 IEEE 8th Southern Power Electronics Conference and 17th Brazilian Power Electronics Conference (SPEC/COBEP)*, 2023, pp. 1–7.