

Deep reinforcement learning-based energy management of hybrid battery systems in electric vehicles

Weiham Li^{a,b,*}, Han Cui^{a,b}, Thomas Nemeth^{a,b}, Jonathan Jansen^{a,b}, Cem Ünlübayır^{a,b}, Zhongbao Wei^e,
Lei Zhang^e, Zhenpo Wang^e, Jiageng Ruan^f, Haifeng Dai^g, Xuezhe Wei^g, Dirk Uwe Sauer^{a,b,c,d}

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

^b*Juelich Aachen Research Alliance, JARA-Energy, Germany*

^c*Institute for Power Generation and Storage Systems (PGS), E.ON ERC, RWTH Aachen University, Germany*

^d*Helmholtz Institute Münster (HI MS), IEK-12, Forschungszentrum Jülich, Germany*

^e*National Engineering Laboratory for Electric Vehicles, School of Mechanical Engineering, Beijing Institute of Technology, Beijing, China*

^f*Faculty of Materials and Manufacturing, Beijing University of Technology, Beijing, China*

^g*National Fuel Cell Vehicle & Powertrain System Research & Engineering Center, School of Automotive Studies, Tongji University, Shanghai, China*

Abstract

In this paper, an energy management strategy based on deep reinforcement learning is proposed for a hybrid battery system, which consists of a high-energy and a high-power battery pack, in electric vehicles. The energy management strategy of the hybrid battery system was developed on the basis of electrical and thermal characterization of the battery cells, aiming at minimizing the energy loss and increasing both the electrical and thermal safety level of the whole system. Especially, we designed a novel reward term to explore the optimal operating range of the high-power pack without imposing a rigid constraint of state of charge. Furthermore, various load profiles were randomly combined to train the deep Q-learning model, which avoided the overfitting problem. The training and validation results showed both the effectiveness and the reliability of the proposed strategy in loss reduction and safety enhancement. The proposed energy management strategy has demonstrated its superiority over the reinforcement learning-based methods in both computation time and energy loss reduction of the hybrid battery system, highlighting the use of such method in future energy management systems.

Keywords: lithium-ion battery, hybrid battery system, reinforcement learning, deep Q-learning, energy management, electric vehicle

1. Introduction

The development of battery electric vehicles (BEVs) has drawn considerable attention in the past decades under the consideration of fossil resource depletion and climate change. With zero driving emission, high

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

Email address: batteries@isea.rwth-aachen.de, weiham.li@isea.rwth-aachen.de (Weiham Li)

Nomenclature

C	heat capacity	List of abbreviations	
C_N	battery nominal capacity	BEV	battery electric vehicles
$C_{1,2}$	battery polarization capacitances	CP	convex programming
E	expected value	DNN	deep neural network
L	mean squared time difference error	DP	dynamic programming
P_t	total power at time t	DQL	deep Q-learning
Q	Q value	ECM	equivalent circuit model
Q^*	optimal Q value	$ECMS$	equivalent consumption minimization strategy
R	total reward	EMS	energy management system
R_0	battery ohmic resistance	FTP	federal test procedure
$R_{1,2}$	battery polarization resistances	HBS	hybrid battery system
T	battery average temperature	HE	high-energy
$V_{1,2}$	battery polarization voltages	HP	high-power
α	weighting factors	LMO	lithium-manganese-oxide
$\dot{Q}$	heat generation rate	LTO	lithium-titanate-oxide
η	coulomb efficiency	MPC	model predictive control
γ	discount factor	NCA	lithium-nickel-cobalt-aluminum-oxide
θ	network weights	$NEDC$	new european driving cycle
ξ	learning rate	OCV	open circuit voltage
a_t	action at time t	QL	Q-learning
h	heat transfer coefficient	$ReLU$	rectified linear unit
r_t	immediate reward at time t	RL	reinforcement learning
s_t	state at time t	SBS	single battery system
t	time	SoC	state of charge
v_t	vehicle velocity at time t	$WLTC$	worldwide harmonized light vehicles test cycles

powertrain efficiency and integration possibility of renewable energies, BEV is one of the most environmentally friendly vehicles compared with the traditional internal combustion engine vehicles. However, challenges are still existing for the design and operation of battery systems in BEVs to further reduce the cost and increase the performance and lifetime. On the one hand, the size of the battery system in BEVs should not be oversized, considering the cost and system efficiency. On the other hand, the requirements of both energy and power capability of the battery system should be satisfied under all possible situations, especially in operating conditions with low temperatures and extreme state of charge (SoC) [1].

In order to overcome the challenges mentioned above, the hybridization of the energy source has been considered widely in the literature. Hybrid energy storage systems usually combine a high energy density storage technology with a high power density storage technology via power electronics. Different storage technologies, such as super-capacitors [2], have been used to cover the requirement of power capability in the hybrid energy storage system. Although super-capacitors show high efficiency, high cycle life and high power density even at low temperatures, several drawbacks, e.g., low energy density and high cost-per-energy, still exist. With the appearance of high power lithium-ion batteries on the market, e.g., batteries with

lithium-titanate-oxide (LTO) anode [3], hybrid battery systems (HBSs) attract more and more attention from industry [1, 4, 5]. Compared with a graphite anode, LTO offers higher cycle lifetime, better thermal stability and higher charging and discharging current without lithium-plating and the formation of a passivating solid electrolyte interface layer due to the high lithium-titanate potential (1.55 V) vs. Li. Compared with the single battery pack topology, HBSs can balance the power and energy demand of the BEVs, which makes the scalability of the entire system possible. As a result, lower volume and less weight can be achieved [4]. However, the effective operation of the HBS relies on an energy management system (EMS), which controls the power distribution between the high-energy (HE) and high-power (HP) battery packs.

A large number of energy management strategies have been proposed in the literature [6] and can be roughly divided into three categories: rule-based, optimization-based, and learning-based methods. The rule-based EMS, e.g., thermostatic strategy [7, 8] and fuzzy logic [9], is usually based on heuristic operation rules and achieved success in the automotive industry because of its simplicity, ease of real-time implementation and high reliability. Nonetheless, the performance of rule-based methods depends significantly on the specific knowledge of an expert as well as on the extensive experimental dataset. The performance of these rules is also far from optimality considering different control objectives. Therefore, rule-based EMS can only be used for specific driving conditions.

The optimization-based methods, e.g., particle swarm optimization [10], equivalent consumption minimization strategy (ECMS) [11–13], dynamic programming (DP) [5, 14, 15] and model predictive control (MPC) [16], are applied to EMSs on the basis of the fixed driving cycles or predicted driving conditions in the future. The overarching goal of this method is the optimization of one or multiple predefined objective values under consideration of system constraints. Yang et al. [11] proposed an ECMS-based EMS to control the power allocation between the fuel cell and the battery for a plug-in hybrid electric bus with adaptive equivalent factors. Through the instantaneous minimization of the cost index, the local optimization problem can be solved. In Ref. [17], the DP-based EMS was applied to achieve optimal control within a hybrid energy storage system. Peng et al. [18] developed a parallel DP-based algorithm based on the matrix calculation as the EMS for a fuel cell and battery hybrid train. However, the future driving information is usually needed in advance, which limits the application of the DP-based EMS to the offline benchmark for the evaluation of other methods. Although the stochastic DP is proven effective in an online implementation [1], an additional Markov transition model for vehicle velocity and power demand is required, which affects the performance of the trained EMS. Another optimization-based method, convex programming (CP), was also explored in Ref. [19, 20] for energy management in hybrid electric vehicles and multi-motor-driven electric vehicles. CP offers a faster calculation due to the convexity with the degradation in optimality compared with DP. To explore the online optimization-based energy management, MPC was adopted in [16] for the design of a strategy with an explicit consciousness of degradation of both battery and fuel cell in hybrid electric vehicles.

State-of-the-art learning-based methods, e.g., reinforcement learning (RL), are becoming one of the most popular methodologies for model-free and real-time energy management [21]. They can learn from the historical experiences and adapt the strategy gradually by maximizing the estimated total reward. The main difference between DP- and RL-based algorithms is that the latter doesn't assume the knowledge of an exact mathematical model of the Markov decision process. With the revolution of the battery management systems with cloud computing and the internet of things, battery relevant data can be measured and transmitted to the cloud seamlessly [22, 23], where learning-based energy management approaches will show significant advantages over the other methods facing with a large amount of operation data. As a typical RL method, Q-learning (QL) is then introduced to solve the energy management problem, in which a Q-table is implemented to store the action values (Q values) of all possible state-action combinations. Biswas et al. [24] proved that the QL-based EMS can achieve a near-global optimal control for the electrified powertrain. Xiong et al. [25] introduced a QL-based EMS for a hybrid energy storage system containing a battery pack and a super-capacitor. Although this EMS can reduce the total energy loss compared with rule-based methods, its performance is dependent on the discretization of both the environmental states and the action space. The training is discounted drastically with the dimension buildup of state and action space, i.e., the strategy suffers from the so-called "curse of dimensionality." Compared with QL, the deep Q-learning (DQL) uses multi-layer neural networks to approximate the Q-matrix, enabling an obvious improvement towards the continuous state space. Hence, the DQL outperforms the QL for solving the optimization problem with multidimensional states [26–30]. Attributed to its superior performance, the DQL has gained great success in various areas such as PC games, autonomous driving, and robotic control [31–34]. Recently, the DQL was employed for the energy management of electric vehicles, achieving satisfying performance concerning the minimization of the fuel economy for plug-in hybrid electric vehicles [26, 35–37] and for hybrid electric vehicles [27, 38]. In Ref. [39], a DQL-based EMS minimized the electricity consumption within a hybrid energy storage system consisting of a battery pack and a super-capacitor.

To the best knowledge of the authors, no efforts have been made to develop a learning-based energy management strategy for HBSs in BEVs. The present work aims to bridge the aforementioned research gap by integrating the following main contributions:

- Development of a DQL-based energy management strategy for HBSs in BEVs with the perspective of energy loss minimization and electrical and thermal safety enhancement.
- Novel design of a reward function for the automatic determination of the optimal operating range of the HP battery pack considering the power demand and power distribution in the HBS.
- Development and parameterization of a coupled electro-thermal battery model with characterization tests to simulate the electrical and thermal dynamics of both HE and HP battery cells.

Additionally, this work extends the existing research with the following secondary contributions:

- Design of a new training scenario to restrain the overfitting problem by combining different driving conditions randomly in each epoch to simulate the uncertainty of the driving conditions.
- Validation of the proposed DQL-based energy management strategy with new driving conditions and comparative study with a QL-based strategy for the same system, highlighting the superiority of the proposed method in both computation and performance.

2. Hybrid battery system model

The energy source of the BEVs on the market is generally composed of a single battery type. Its energy and power capability relies on the specific cell type, which usually leads to an oversizing of the battery system. Using a hybrid system with different battery types offers the possibility to fulfill the power and energy requirements of BEVs more efficiently.

2.1. Electric vehicle and HBS modeling

The BEV model with an HBS was developed in our previous work, which includes two electric motors, a high voltage DC-DC converter, two traction inverters, and a battery system. Each axle is propelled by one electric motor, resulting in an all-wheel-drive configuration. Considering the balance between the system flexibility and cost, we implement one DC-DC converter in the HBS, as shown in Fig. 1. The HE battery pack is connected directly to the DC-link. The modeled DC-DC converter between the HP battery pack and the DC-link is Brusa BDC546, whose energy loss is modeled with a second-order polynomial based on input current [17] as follows,

$$Q_{DC-DC} = 1.56 \times 10^{-2} I_{DC-DC}^2 - 1.44 I_{DC-DC} + 388.90 \quad (1)$$

where Q_{DC-DC} and I_{DC-DC} are the energy loss and current of the DC-DC converter, respectively. The readers are referred to [40] for modeling details of the BEV.

2.2. Battery modeling and parameterization

In this work, both HE and HP battery packs are constructed with battery cells of a single battery type but different electrode material compositions. The HE battery pack is supposed to support the base load of the BEV, while the HP battery pack serves as an additional power source to fulfill the dynamic load. The HE battery pack and HP battery pack are composed of 4.9 Ah HE cells and 2.9 Ah HP cells, respectively. The HE cells with lithium-nickel-cobalt-aluminum-oxide (NCA) and graphite as electrode materials have a very high energy density but limited power capability. In contrast, the HP cells with lithium-manganese-oxide (LMO) and LTO as the cathode and anode materials are characterized by the high power density but low

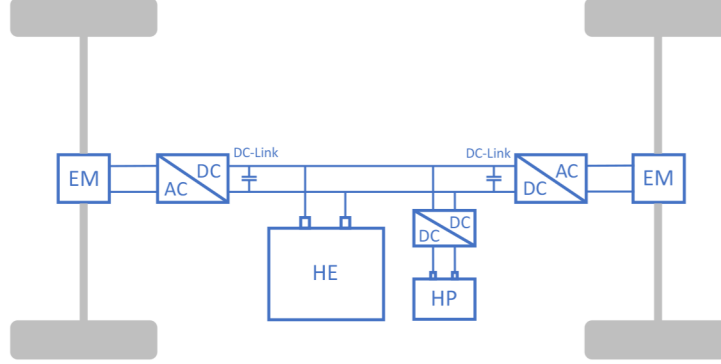


Figure 1: The topology of the HBS for the BEV.

energy density. The specifications of the battery cells and the configuration of the HBS are adopted from [5] and summarized in Table 1.

	HE cell	HP cell
Chemistry (Anode / Cathode)	C / NCA	LTO / LMO
Cell nominal capacity	4.9 Ah	2.9 Ah
Cell nominal voltage	3.6 V	2.4 V
Cell voltage limits (min / max)	2.5 V / 4.2 V	1.5 V / 2.9 V
Cell current limits (DCH / CHA)	2 C / 1 C	70 C / 70 C
Energy density	250 Wh/kg	45 Wh/kg
Power density (10 s, DCH / CHA)	1.3 kW/kg / 0.5 kW/kg	3.2 kW/kg / 3.2 kW/kg
Cell weight	69 g	150 g
Pack configuration	90s 16p	90s 3p
Pack energy	25.4 kWh	1.9 kWh

Table 1: Specifications of the HBS.

2.2.1. Equivalent circuit model

With the assumption that the battery cells in each pack are the same and are working in the same conditions, two equivalent circuit models (ECMs) are developed and parameterized with experiments conducted on HE and HP battery cells, respectively. Compared with electrochemical models [41, 42], ECMs [43–46] offer the benefits in computing speed due to the simple model structure, which further reduces the training time and improves the training efficiency of the proposed DQL-based energy management strategy. Considering the trade-off between model accuracy and computational burden, we chose an extended Thevenin model with two RC pairs to simulate the dynamics of battery cells. The model contains the open-circuit voltage (OCV), which has a nonlinear relationship with the state of charge (SoC), as the voltage source, ohmic resistance R_0 , polarization resistances $R_{1,2}$ and polarization capacitances $C_{1,2}$, as shown in Fig. 2.

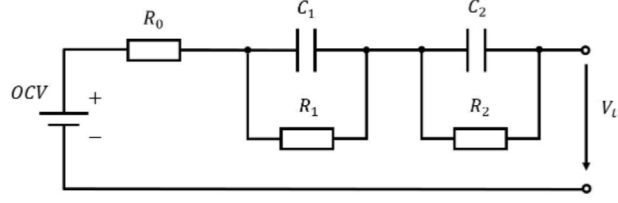


Figure 2: Extended Thevenin model of the battery cells.

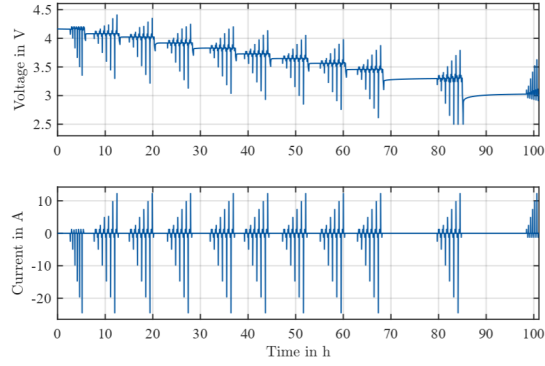


Figure 3: Pulse tests of the HE cell at 25 °C.

The model is described as follows:

$$\dot{SoC}(t) = \frac{\eta}{C_N} I(t) \quad (2)$$

$$\dot{V}_1(t) = -\frac{1}{R_1 C_1} \cdot V_1(t) + \frac{1}{C_1} I(t) \quad (3)$$

$$\dot{V}_2(t) = -\frac{1}{R_2 C_2} \cdot V_2(t) + \frac{1}{C_2} I(t) \quad (4)$$

$$V_t(t) = OCV(SoC(t)) + V_1(t) + V_2(t) + R_0 I(t) \quad (5)$$

where $I(t)$ and $V_t(t)$ are the current and terminal voltage, respectively, η is the coulomb efficiency, C_N represents the nominal capacity of the battery cell, $V_1(t)$ and $V_2(t)$ are voltages over R_1 and R_2 , respectively.

To get the accurate parameters of the ECM, we carried out a capacity test at first to determine the cell capacity. The OCV was determined by measuring the terminal voltage of a fully relaxed cell at different SoCs. Then pulse tests were performed under different conditions from 100% to 0% SoC with 10% intervals at five different C-rates and temperatures, as shown in Fig. 3. The detailed specifications of the pulse tests are summarized in Table 2.