IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received May 5, 2021, accepted May 13, 2021, date of publication May 17, 2021, date of current version May 26, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3081184

Economic Adaptive Cruise Control for Electric Vehicles Based on ADHDP in a Car-Following Scenario

XIYAN CHEN[1,*], JIAN YANG[2], CHUNJIE ZHAI[1,*], (Member, IEEE), JIEDONG LOU[1], AND CHENGGANG YAN[1]

[1]School of Automation, Hangzhou Dianzi University, Hangzhou 310018, China
[2]HDU-ITMO Joint Institute, Hangzhou Dianzi University, Hangzhou 310018, China

Corresponding author: Chunjie Zhai (chunjiezhai2@gmail.com)

*Xiyan Chen and Chunjie Zhai contributed equally to this work.

This work was supported in part by the Natural Science Foundation of Zhejiang Province under Grant LQ21F030013; in part by the National Natural Science Foundation of China under Grant 61931008, Grant 61671196, and Grant 51906053; in part by the National Program on Key Basic Research Project under Grant 2017YFC0820605; and in part by the Research Startup Foundation of Hangzhou Dianzi University under Grant KYS065620050.

ABSTRACT In order to ensure vehicle safety, enhance riding comfort, extend the battery life of electric vehicles (EVs), and improve the energy economy, an ADHDP-based economic adaptive cruise control (Eco-ACC) strategy for EVs in car-following scenarios is proposed in this paper. First, the longitudinal dynamics of EVs is modeled, and the control objectives are presented; then, the actor-critic structure of ADHDP is introduced, and the policy iteration formulas of the critic and actor networks in the ADHDP framework are given; finally, after the state variables, control variables, unity function and value function are determined, the ADHDP-based Eco-ACC strategy for EVs is designed. Extensive simulation results under different driving cycles show that the proposed Eco-ACC strategy can not only ensure vehicle safety, improve riding comfort and reduce energy consumption, but also significantly reduce the battery capacity loss and extend the battery life compared with the benchmark algorithm. In addition, the proposed Eco-ACC strategy is model-free and real-time, and can be robust in different car-following scenarios.

INDEX TERMS Economic adaptive cruise control, ADHDP, vehicle safety, battery lifespan, energy economy.

I. INTRODUCTION

As a current energy-intensive industry, the transportation sector is facing the pressure of energy saving and emission reduction [1]–[3]. To address such a challenge, increasing effort and massive energy have been devoted to the research of energy-efficient vehicles. Alternative energy sources and hybrid/electric vehicle technologies have been widely investigated, and they have shown great potential for significant energy savings and emission reductions [4]. In addition, intelligent driving technology can also be used to increase the intelligence of energy-efficient vehicles, thus further improve energy economy and reduce driving cost [5].

Advanced driver assistant system (ADAS) is the initial development stage of intelligent driving technology. It can use various on-board sensors to obtain relevant data and realize automatic control of the vehicle [6]. As an ADAS system, adaptive cruise control (ACC), which is extended from the cruise control, can automatically adjust the speed of the controlled vehicle to maintain a desired safe distance [7], [8], which is conducive to improving traffic flow, reducing traffic accidents and providing comfort driving experience [9]. The commonly used methods for the design of ACC strategies include model predictive control (MPC) [10], [11], sliding mode control [12] and PID control [13]. Even though the traditional ACC system can achieve vehicle safety and reduce energy consumption, its energy saving effect is not significant, especially for passenger cars with small frontal area. In order to further improve the energy economy, the economic adaptive cruise control (Eco-ACC), which combines ACC and ecological driving technology [14], has become a research topic [4], [15].

The Eco-ACC strategies are usually designed to maximize energy economy under the premise of ensuring vehicle

The associate editor coordinating the review of this manuscript and approving it for publication was Mohamad Afendee Mohamed.

safety. To minimize fuel consumption of fossil-based road vehicles in car-following scenarios, Li et al presented a periodic servo-loop longitudinal control algorithm for an ACC system, where the fuel-saving mechanism of pulse-and-glide operation was first discussed [16]. Jia *et al.* proposed an energy-optimal ACC for electric vehicles (EVs) based on MPC, which plans the host car's speed trajectory in real time for higher energy efficiency by taking look-ahead traffic information and road conditions into consideration [17]. To minimize fuel consumption of heavy duty vehicles in a highway scenario, Borek *et al.* designed an economic optimal control strategy that combines dynamic programming (DP) and MPC, in which the DP method is used to obtain the global optimal velocity trajectory while the MPC is employed to track the reference speed and optimize the fuel consumption [18].

Due to time-varying parameters, limited detection capabilities and external disturbances, car following systems often have multiple uncertainties, making it difficult to accurately model. The above Eco-ACC strategies based on the accurate models often have poor control performance or even instability in real car following systems. To deal with the speed disturbance of the preceding vehicle, Moser *et al.* proposed a flexible spacing ACC control method based on stochastic MPC, which uses a conditional linear Gaussian model to estimate the probability distribution of the future velocity of the preceding vehicle [19]. Sakhdari and Azad proposed an adaptive tube-based nonlinear MPC method for economic autonomous cruise control of plug-in hybrid electric vehicles, which can improve the economy and tracking of ACC systems while maintaining robustness to disturbances and modeling errors [9]. To reduce the energy consumption, Lee *et al.* presented a novel model-based reinforcement learning algorithm for eco-driving control of EVs, where the domain knowledge of vehicle dynamics and the powertrain system is utilized for the reinforcement learning process while model-free characteristics are maintained by updating the approximation model using experience replay [20]. In addition, Li and Görges put forward an model-free Eco-ACC strategy based on reinforcement learning for fuel vehicles, which can improve fuel economy and enhance vehicle safety [21].

The EV has been considered as one of the most energy-efficient vehicles due to their high efficiency and zero emission. However, the model-free Eco-ACC strategies oriented towards EVs are very few, and the existing model-free strategies for fuel vehicles cannot be applied to the EVs. In addition, according to [22], the battery life directly affects the full life cycle cost of EVs, and is one of the key factors that determine the market share of EVs. Nevertheless, the existing Eco-ACC strategies for EVs pay more attention to vehicle safety, comfort and energy economy, and hardly considers the battery life of EVs. It would be very economic if the battery lifespan of the EVs can be significantly prolonged while improving the energy economy, which will promote the popularization of EVs.

This paper aims to propose an Eco-ACC strategy for EVs based on ADHDP in a car-following scenario, which does not rely on the domain knowledge of vehicle dynamics and can obtain the approximate optimal control effort under model-free conditions. Compared with the existing works, the main contributions of this paper are as follows: (1) a model-free and real-time Eco-ACC strategy based on ADHDP is proposed for EVs in a car-following scenario to ensure vehicle safety, enhance riding comfort, prolong the battery life of EVs and improve energy economy; (2) different from the existing model-free Eco-ACC strategies, the concept of equivalent spacing deviation is defined using the band-stop function with compensating factors, which can constrain the distance between adjacent vehicles within the allowable range; (3) extensive simulations demonstrate the robustness of the proposed Eco-ACC strategy in different car-following scenarios.

The remainder of this paper is organized as follows: Section II presents the longitudinal dynamics model of EVs and the control objectives, followed by the principle and policy iterations of ADHDP in Section III. In Section IV, the ADHDP-based Eco-ACC strategy for the EVs is presented. In Section V, the simulation results under different driving cycles are discussed, followed by the conclusion of this paper.

II. MODELING AND CONTROL OBJECTIVES

This paper attempts to develop an Eco-ACC strategy for an EV to ensure vehicle safety, prolong battery lifespan and decrease energy consumption in a car-following scenario. The car-following system studied in this paper can be seen in FIGURE 1, where the EV can use the vehicle radar to detect the distance and relative speed to the vehicle ahead. In addition, the system architecture of EVs is presented in FIGURE 2, where the motor is connected to the battery pack through the DC/AC inverter, and the transmission shaft to the motor through the gear box, of which the gear ratio is assumed to be fixed. This section will present the longitudinal dynamics model of EVs and the control objectives of the proposed Eco-ACC strategy.

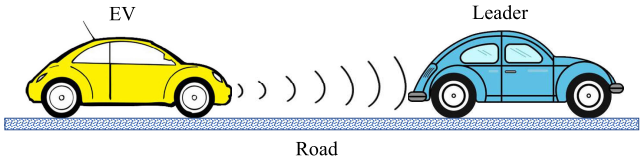


FIGURE 1. The car-following scenario.

A. LONGITUDINAL DYNAMICS MODEL OF EVS

In the existing studies, different models are used to describe the longitudinal dynamics of a vehicle, such as the second-order models [2], [17], [18], and the third-order ones [4], [7], [9]. In this paper, the second-order model will be employed to describe longitudinal dynamics of the EV, which

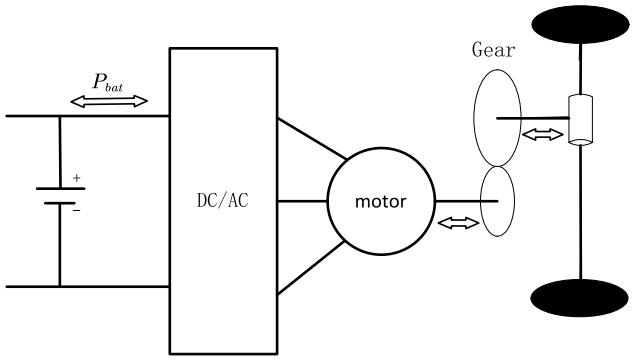


FIGURE 2. The system architecture of EVs.

can be described as

$$\begin{cases} \dot{s}(t) = v(t) \\ \dot{v}(t) = u(t) \end{cases} \tag{1}$$

where $s(t)$, $v(t)$ and $u(t)$ are the position, velocity, control input of the EV at time t, respectively.

Furthermore, we can have

$$\delta m u(t) = \frac{\eta_t}{R} T_w(t) - F_b(t) - F_r(t), \tag{2}$$

where δ, m, η_t, $T_w(t)$, $F_b(t)$ and $F_r(t)$ are the rotational inertia coefficient, mass, mechanical efficiency, traction, braking force and integrated resistance force of the EV.

According to [4], the integrated resistance force $F_r(t)$ can be given as

$$F_r(t) = mg(\mu cos(\theta(s(t))) + sin(\theta(s(t)))) + \frac{C_d \Phi(d(t)) \rho A_v v^2(t)}{2}, \tag{3}$$

where C_d, $\Phi(d(t))$, A_v, μ and $\theta(s(t))$ are the drag coefficient for solo driving, the normalized drag coefficient related to vehicle spacing $d(t)$, frontal area, rolling coefficient and road slope related to position, respectively; ρ and g are the air density and gravitational acceleration, respectively. And the spacing $d(t)$ between the EV and its preceding vehicle can be calculated through

$$d(t) = s(t) - s_p(t) - L, \tag{4}$$

where L is the length of the EV, and $s_p(t)$ is the position of the preceding vehicle.

The torque $T_m(t)$ and rotational speed $\omega_m(t)$ at the motor can be obtained as

$$\begin{cases} T_m(t) = T_w(t)/G_r \\ \omega_m(t) = v(t) * G_r/R, \end{cases} \tag{5}$$

where R and G_r are the tire radius and the fixed gear ratio, respectively.

Then the input power of motor-inverter $P_{bat}(t)$ can be given as

$$P_{bat}(t) = \begin{cases} T_m(t)\omega_m(t)/\eta_m(t), & \text{if } T_m(t) \geq 0 \\ T_m(t)\omega_m(t)\eta_m(t), & \text{otherwise.} \end{cases} \tag{6}$$

where $\eta_m(t)$ can be calculated through

$$\eta_m(t) = f_m(\omega_m(t), T_m(t)). \tag{7}$$

Note that f_m is the power transfer efficiency of the motor-inverter, which can be experimentally obtained from a bench test [23].

According to [5], the maximum motor torque of the EV can be obtained as

$$T_{m,max}(t) = \begin{cases} 198, & 0 \leq \omega_m(t) < 244 \\ -0.1094\omega_m(t) + 224.7, & 244 \leq \omega_m(t) < 308 \\ 18470\omega_m^{-0.7389}(t) - 74.78, & \omega_m(t) \geq 308, \end{cases} \tag{8}$$

where the unit of the rotational speed is rad/s.

Due to the physical limitations, the torque at the motor is constrained as

$$-T_{m,max}(t) \leq T_m(t) \leq T_{m,max}(t). \tag{9}$$

According to [5], the braking torque, which is required to decelerate vehicles, can be provided by the mechanical braking system and the regenerative braking system. Only if the required braking torque is larger than the maximum braking torque that can be provided by the motor, the rest of the required braking torque will be offered by the mechanical braking system.

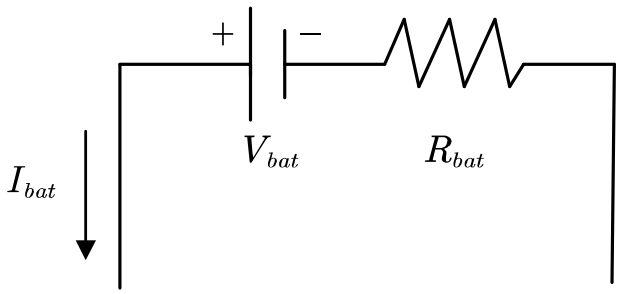


FIGURE 3. The Rint model of the battery pack.

In this paper, the battery pack is assembled by lithium cells, and the behaviors of the battery pack are represented using the control-oriented *Rint* model in FIGURE 3. According to the energy conversion principle and the Kirchhoff's voltage and current laws, the relationship among the battery voltage $V_{bat}(t)$, the battery current $I_{bat}(t)$ and the battery output power $P_{bat}(t)$ is given as

$$P_{bat}(t) = V_{bat}(t)I_{bat}(t) - I_{bat}^2(t)R_{bat}(t), \tag{10}$$

where, according to [5], the battery resistance $R_{bat}(t)$ is time-varying instead of being a constant.

B. CONTROL OBJECTIVES

This paper aims to propose a model-free Eco-ACC strategy in order to achieve vehicle safety, extend battery life, and improve energy economy, which can be defined as follows:

1) VEHICLE SAFETY

To ensure vehicle safety, according to [24], the vehicle spacing should be limited as

$$d_{min}(t) \le d(t) \le d_{max}(t), \tag{11}$$

where $d_{min}(t)$ and $d_{max}(t)$ are the allowed minimum and maximum spacings, respectively, and they can be calculated through

$$\begin{cases} d_{min}(t) = 2 + 0.5 \cdot v(t) + 0.0625 \cdot v^2(t) \\ d_{max}(t) = 10 + v(t) + 0.0825 \cdot v^2(t). \end{cases} \tag{12}$$

2) RIDING COMFORT

To ensure riding comfort, similar to [4], the control input of the EV should be limited as

$$a_{min} \le u(t) \le a_{max}, \tag{13}$$

where a_{min} and a_{max} are the allowed minimum and maximum accelerations, respectively. It should be noted that a_{min} and a_{max} should be properly determined without violating the physical limits of the power and braking systems. In this paper, a_{min} and a_{max} are set as $a_{min} = -2$ m/s^2 and $a_{max} = 2$ m/s^2, respectively.

3) LONGER BATTERY LIFE

According to [25], lowering $I^2_{bat}(t)$ can decrease the batteries' duty, and further extend the battery life. Therefore, in order to prolong the battery lifespan, it is necessary to reduce

$$\int_{t_0}^{T_{cyc}} I^2_{bat}(t)dt, \tag{14}$$

where t_0 and T_{cyc} are the initial and terminal time of a journey.

4) ENERGY ECONOMY

The power of the EV comes from the battery pack, to improve energy economy, it is helpful to decrease

$$\int_{t_0}^{T_{cyc}} V_{bat}(t) I_{bat}(t)dt. \tag{15}$$

III. ACTION-DEPENDENT ADAPTIVE DYNAMIC PROGRAMMING

Reinforcement learning is an effective method, which can make an agent learn how to make optimal decisions through interactions with the dynamic environment [26], and adaptive dynamic programming is an important branch of reinforcement learning in the field of control. As one basic framework of adaptive dynamic programming, the ADHDP cannot rely on the information of the object and environment, and will be used to design the model-free Eco-ACC strategy. In this section, the actor-critic structure and its policy iterations of the ADHDP will be presented in detail.

A. THE ACTOR-CRITIC STRUCTURE

A general nonlinear discrete-time system can be given as

$$\boldsymbol{x}(k+1) = f(\boldsymbol{x}(k), \boldsymbol{u}(k)), \quad k = 0, 1, 2, \ldots \tag{16}$$

where $\boldsymbol{x} \in \mathbb{R}^m$ and $\boldsymbol{u} \in \mathbb{R}^n$ are the state and control vectors, respectively.

The value function of the above discrete-time system at time instant k can be obtained as

$$\begin{aligned} J(x(k)) &= \sum_{i=k}^{\infty} \gamma^{i-k} U(\boldsymbol{x}(i), \boldsymbol{u}(i)) \\ &= U(\boldsymbol{x}(k), \boldsymbol{u}(k)) + \gamma \sum_{i=k+1}^{\infty} \gamma^{i-(k+1)} U(\boldsymbol{x}(i), \boldsymbol{u}(i)) \\ &= U(\boldsymbol{x}(k), \boldsymbol{u}(k)) + \gamma J(\boldsymbol{x}(k+1)), \end{aligned} \tag{17}$$

where $U(\boldsymbol{x}(k), \boldsymbol{u}(k))$ is the unity function at time instant k, and $\gamma (0 < \gamma \le 1)$ indicates the discount factor.

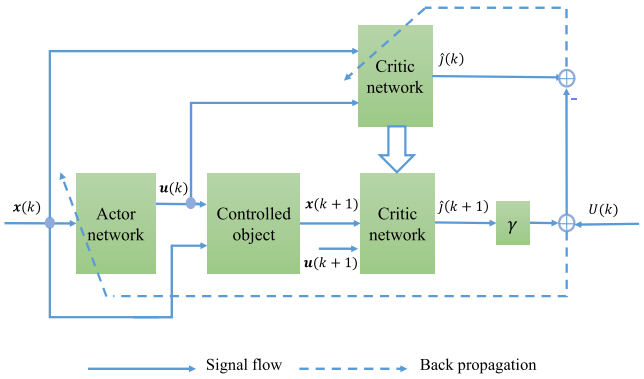


FIGURE 4. The actor-critic structure.

According to the Bellman optimality principle, we can have

$$J^*(\boldsymbol{x}(k)) = U(\boldsymbol{x}(k), \boldsymbol{u}^*(k)) + \gamma J^*(\boldsymbol{x}(k+1)), \tag{18}$$

where $J^*(\boldsymbol{x}(k))$ is the optimal value function for the optimal policy $\boldsymbol{u}^*(k)$ at time instant k.

And the optimal policy $\boldsymbol{u}^*(k)$ in (18) can be obtained through

$$\boldsymbol{u}^*(k) = arg \min_{\boldsymbol{u}(k)} [U(\boldsymbol{x}(k), \boldsymbol{u}(k)) + \gamma J^*(\boldsymbol{x}(k+1))]. \tag{19}$$

To achieve the optimal policy, in the ADHDP framework, the actor-critic structure in FIGURE 4 is usually used, where the actor and the critic are usually constructed by the BP neural networks. To solve the optimal problem in equation (19), the critic network and the actor network are used to approximate the optimal value function $J^*(\boldsymbol{x}(k))$ and the optimal control policy $u^*(k)$, respectively [21], [27].

B. POLICY ITERATION

After the actor-critic structure is introduced, the policy iteration of critic network and actor network for solving the Bellman optimality equation in (18) will be detailed.

1) CRITIC NETWORK

As shown in FIGURE 5, the inputs of the critic network are the m-dimensional state vector $\boldsymbol{x}(k)$ and the n-dimensional control vector $\boldsymbol{u}(k)$, and the output is an estimate of $\hat{J}^*(\boldsymbol{x}(k))$. In addition, the hyperbolic tangent transfer function $\phi(x) = \frac{1-e^{-x}}{1+e^{-x}}$ is used in the critic network.

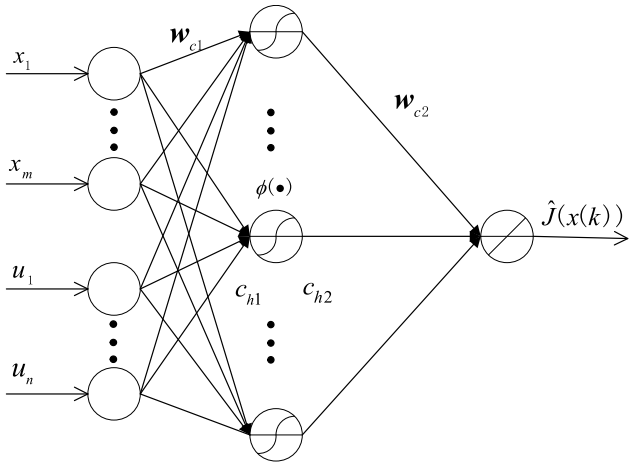


FIGURE 5. The structure of the critic network.

The input vector of the critic network at time instant k can be defined as

$$\boldsymbol{C}(k) = [x_1(k), \ldots, x_m(k), u_1(k), \ldots, u_n(k)] \tag{20}$$

The inputs and outputs of the hidden layer in the critic network at time instant k can be calculated through

$$\begin{cases} c_{h1}^{(j)}(k) = \sum_{i=1}^{n+m} \boldsymbol{C}^{(i)}(k) \cdot \boldsymbol{w}_{c1}^{(i,j)}(k), \\ c_{h2}^{(j)}(k) = \dfrac{1 - e^{-c_{h1}^{(j)}(k)}}{1 + e^{-c_{h1}^{(j)}(k)}}, \end{cases} \tag{21}$$

where $c_{h1}^{(j)}(k)$ and $c_{h2}^{(j)}(k)$ are the input and output of the jth node in the hidden layer of the critic network, respectively; $\boldsymbol{C}^{(i)}(k)$ is the ith dimensional component of the input vector $\boldsymbol{C}(k)$; $\boldsymbol{w}_{c1}^{(i,j)}(k)$ is the weight from the ith node of the input layer to the jth node of the hidden layer in the critic network.

Then, the output of the critic network can be given as

$$\hat{J}(k) = \sum_{j=1}^{N_c} c_{h2}^{(j)}(k) \cdot \boldsymbol{w}_{c2}^{(j)}(k), \tag{22}$$

where $\boldsymbol{w}_{c2}^{(j)}(k)$ is the weight from the jth node of the hidden layer to the only output node of the output layer, and N_c is the number of nodes in the hidden layer.

To satisfy the Bellman optimality equation in equation (18), the weights $\boldsymbol{w}_{c1}$ and $\boldsymbol{w}_{c2}$ in the critic network should be updated by minimizing the error $\| E_c \| = \sum_k E_c(k)$, in which $E_c(k)$ can be calculated through

$$E_c(k) = \frac{1}{2} e_c^2(k) = \frac{1}{2} (\hat{J}(k) - U(k) - \gamma \hat{J}(k+1))^2. \tag{23}$$

And a gradient descent adaptation algorithm can be used to update the weights through

$$\begin{cases} \boldsymbol{w}_{c1}^{p+1}(k+1) = \boldsymbol{w}_{c1}^{p}(k) + \Delta \boldsymbol{w}_{c1}^{p}(k) \\ \boldsymbol{w}_{c2}^{p+1}(k+1) = \boldsymbol{w}_{c2}^{p}(k) + \Delta \boldsymbol{w}_{c2}^{p}(k), \end{cases} \tag{24}$$

where p is the iteration number.

Furthermore, $\Delta \boldsymbol{w}_{c1}(k)$ and $\Delta \boldsymbol{w}_{c2}(k)$ can be obtained as

$$\Delta \boldsymbol{w}_{c2}(k) = -\eta_c(k) \cdot e_c(k) \cdot c_{h2}^T(k) \tag{25}$$

and

$$\begin{aligned} \Delta \boldsymbol{w}_{c1}(k) = -\frac{1}{2} \cdot \eta_c(k) \cdot e_c(k) \cdot \boldsymbol{C}^T(k) \\ \times \{ \boldsymbol{w}_{c2}^T(k) \otimes [1 - c_{h2}(k) \otimes c_{h2}(k)] \}, \end{aligned} \tag{26}$$

where $\eta_c(k)$ is the learning rate of the critic network.

2) ACTION NETWORK

As shown in FIGURE 6, the inputs of the actor network are the m-dimensional state vector $\boldsymbol{x}(k)$, and the output is the n-dimensional control vector $\boldsymbol{u}(k)$. Furthermore, the hyperbolic tangent transfer function $\phi(x) = \frac{1-e^{-x}}{1+e^{-x}}$ is also used in the actor network.

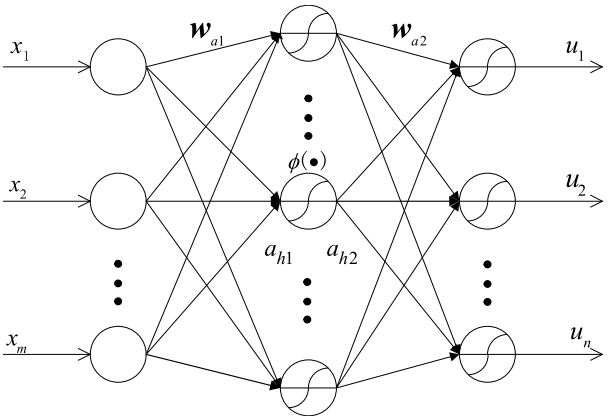


FIGURE 6. The structure of the actor network.

The inputs and outputs of the hidden layer in the actor network at time instant k can be calculated through

$$\begin{cases} a_{h1}^{(j)}(k) = \sum_{i=1}^{m} x_i(k) \cdot \boldsymbol{w}_{a1}^{(i,j)}(k), \\ a_{h2}^{(j)}(k) = \dfrac{1 - e^{-a_{h1}^{(j)}(k)}}{1 + e^{-a_{h1}^{(j)}(k)}}, \end{cases} \tag{27}$$

where $a_{h1}^{(j)}(k)$ and $a_{h2}^{(j)}(k)$ are the input and output of the jth node in the hidden layer of the actor network, respectively; $x_i(k)$ is the ith dimensional component of the state vector $x(k)$; $\boldsymbol{w}_{a1}^{(i,j)}(k)$ is the weight from the ith node of the input layer to the jth node of the hidden layer in the actor network.

Then, the output of the actor network can be given as

$$\begin{cases} u_q(k) = \dfrac{1 - e^{-\mu_q(k)}}{1 + e^{-\mu_q(k)}}, \\ \mu_q(k) = \sum_{j=1}^{N_a} a_{h2}^{(j)}(k) \cdot \boldsymbol{w}_{a2}^{(j,q)}(k) \end{cases} \tag{28}$$

where $\boldsymbol{w}_{a2}^{(i,j)}(k)$ is the weight from the ith node of the hidden layer to the jth node of the output layer in the critic network, and N_a is the number of nodes in the hidden layer.

In the ADHDP, the actor network generates the optimal policy $u^*(k)$ by minimizing the $E_a = \hat{J}(k)$. And the gradient descent adaptation algorithm can be used to update the weights through

$$\begin{cases} \boldsymbol{w}_{a1}^{p+1}(k+1) = \boldsymbol{w}_{a1}^{p}(k) + \Delta \boldsymbol{w}_{a1}^{p}(k) \\ \boldsymbol{w}_{a2}^{p+1}(k+1) = \boldsymbol{w}_{a2}^{p}(k) + \Delta \boldsymbol{w}_{a2}^{p}(k), \end{cases} \tag{29}$$

where p is the iteration number.

Furthermore, $\Delta \boldsymbol{w}_{a1}(k)$ and $\Delta \boldsymbol{w}_{a2}(k)$ can be obtained as

$$\begin{aligned} \Delta \boldsymbol{w}_{a2}(k) &= -\frac{1}{2}\eta_a(k) \cdot [1 - u(k) \otimes u(k)] \cdot a_{h2}^T(k) \times \frac{\partial \hat{J}(k)}{\partial u(k)} \\ &= -\frac{1}{4}\eta_a(k) \cdot [1 - u(k) \otimes u(k)] \cdot a_{h2}^T(k) \times \{\boldsymbol{w}_{c2}^T(k) \\ &\quad \otimes [1 - c_{h2}(k) \otimes c_{h2}(k)]\} \times \boldsymbol{w}_{c1u}^T(k) \end{aligned} \tag{30}$$

and

$$\begin{aligned} \Delta \boldsymbol{w}_{a1}(k) = &-\frac{1}{8}\eta_a(k) \cdot x^T(k) \times [1 - u(k) \otimes u(k)] \\ &\times [(\{\boldsymbol{w}_{c2}^T(k) \otimes [1 - c_{h2}(k) \otimes c_{h2}(k)]\} \times \boldsymbol{w}_{c1u}^T(k) \\ &\times \boldsymbol{w}_{a2}^T(k)) \otimes [1 - a_{h2}(k) \otimes a_{h2}(k)]], \end{aligned} \tag{31}$$

where $\eta_a(k)$ is the learning rate of the actor network; $\boldsymbol{w}_{\mathrm{c1u}} = \boldsymbol{w}_{\mathrm{c1}}(m+1 : end, :)$ and $\boldsymbol{w}_{\mathrm{c1}}(m+1 : end, :)$ is a commonly used matrix representation in MATLAB, representing the last $end - m$ rows of matrix $\boldsymbol{w}_{\mathrm{c1}}$, that is, part of the weights corresponding to the control variables.

IV. THE ADHDP-BASED ECO-ACC STRATEGY

In this section, firstly, the state and control variables in the ADHDP framework are determined; then, the unity function and the corresponding value function are designed; finally, the algorithm of the ADHDP-based Eco-ACC strategy is given.

A. BAND-STOP FUNCTION

To ensure the spacing constraint defined in equation (11), same as [4] and [28], the band-stop function with compensating factors will be used in the proposed Eco-ACC strategy, which can be described as

$$BSF(z|\hat{\alpha}, \hat{\beta}, \hat{n}, z_l, z_u, cf) = \left(\frac{e^{-\hat{\alpha}(z - z_l - cf)} + e^{\hat{\alpha}(z - z_u + cf)}}{\hat{\beta}}\right)^{\hat{n}}, \tag{32}$$

where $\hat{\alpha} > 0$, $\hat{\beta} \geq 1$ and $\hat{n} \in N^+$. $z_l,\ z_u \in \mathbb{R}^+$ are the lower and upper limits of the band $[z_l, z_u]$, respectively. In addition, cf is the compensating factor.

B. STATE AND CONTROL VARIABLES

According to equation (11), the spacing is required to be within dynamic ranges rather than follow a desired value. Here, the equivalent spacing deviation $\Delta d(k)$ between the EV and its preceding vehicle is defined as

$$\Delta d(k) = \begin{cases} BSF(d(k)|\alpha, \beta, \hat{n}, d_{min}(k), d_{max}(k), cf) \\ \text{if } d(k) > \dfrac{d_{min}(k) + d_{max}(k)}{2} \\ -BSF(d(k)|\alpha, \beta, \hat{n}, d_{min}(k), d_{max}(k), cf) \\ \text{if } d(k) \leq \dfrac{d_{min}(k) + d_{max}(k)}{2}, \end{cases} \tag{33}$$

where the parameters α, β, $\hat{n}$ and cf should be set properly.

In order to ensure that the EV can track the preceding vehicle within the allowed spacing range in equation (11), the speed deviation $\Delta v(k)$ and the equivalent spacing deviation $\Delta d(k)$ are selected as the state variables in the ADHDP framework, i.e.,

$$\boldsymbol{x}(k) = [\Delta v(k), \Delta d(k)]^T, \tag{34}$$

where $\Delta v(k) = v_p(k) - v(k)$, and $v_p(k)$ denotes the speed of the preceding vehicle.

The acceleration is selected as the control variable in the ADHDP, i.e., $\boldsymbol{u}(k) = a(k)$. It should be noted that, to ensure the acceleration constraint in equation (13), the control variable should be obtained by linearly mapping the output of the actor network to the internal $[a_{min}, a_{max}]$.

C. UNITY FUNCTION AND VALUE FUNCTION

In order to achieve the aforementioned control objectives in Section II, the unity function in the ADHDP framework can be given as

$$U(\boldsymbol{x}(k), \boldsymbol{u}(k)) = \alpha_1 L_1(k) + \alpha_2 L_2(k) + \alpha_3 L_3(k), \tag{35}$$

where α_1, α_2 and α_3 are the weight coefficients, and L_1, L_2 and L_3 can be described as

$$\begin{cases} L_1(k) = BSF(d(k)|\alpha, \beta, \hat{n}, d_{min}(k), d_{max}(k), cf) \\ L_2(k) = I_{bat}^2(k) \\ L_3(k) = V_{bat}(k) I_{bat}(k). \end{cases} \tag{36}$$

In the equation (35), $L_1(k)$ aims to enforce the spacing to stay in the allowed range $[d_{min}(k), d_{max}(k)]$; $L_2(k)$ is designed to decrease the batteries' duty, and further to extend the battery lifespan; $L_3(k)$ is used to reduce the energy consumption, and thus to improve the energy economy.

After the unity function is obtained, the value function in the proposed Eco-ACC strategy can be given as

$$J(\boldsymbol{x}(k)) = \sum_{i=k}^{\infty} \gamma^{i-k} U(\boldsymbol{x}(i), \boldsymbol{u}(i)). \tag{37}$$

D. THE ALGORITHM OF THE MODEL-FREE ECO-ACC STRATEGY

Once the state vector, the control vector, the unity function and the value function are determined, for the sake of understanding, the algorithm of the proposed ADHDP-based Eco-ACC strategy is given in **Algorithm 1**.

Algorithm 1 The Algorithm of the ADHDP-Based Eco-ACC Strategy

Sampling interval: $\Delta t = 0.1$ s;
State vector: $\boldsymbol{x}(k) = [\Delta d(k), \Delta v(k)]^T$;
Control vector: $\boldsymbol{u}(k) = a(k)$;
Randomly initialize weights of the critic and actor networks;
Iteration number of the critic network: i_c;
Iteration number of the actor network: i_a;
Allowed maximum iteration number of the critic network: N_{cm};
Allowed maximum iteration number of the actor network: N_{am};
Allowed error of critic network: T_c;
Allowed error of actor network: T_a;
for $k = 1$ to $10 * T_{cyc}$ **do**
  Calculate $\Delta d(k)$ and $\Delta v(k)$ through equations (11), (33)-(34);
  Calculate $E_c(k)$ through equations (20)-(23), set $i_c = 0$;
  while $E_c(k) > T_c$ & $i_c < N_{cm}$ **do**
    Update weights $\omega_c^{(i_c+1)} = \omega_c^{(i_c)} + \Delta\omega_c^{(i_c)}$ with equations (24)-(26);
    Update $\hat{J}(k)$ through equations (20)-(22);
    Update $E_c(k)$ through equations (20)-(23);
    Set $i_c = i_c + 1$;
  end while
  Calculate $\hat{J}(k)$ through equations (20)-(22), set $\Delta_a = \hat{J}(k)$, $\delta_a = 0$ and $i_a = 0$;
  while $|\Delta_a - \delta_a| > T_a$ & $i_a < N_{am}$ **do**
    $\delta_a = \Delta_a$;
    Update weights $\omega_a^{(i_a+1)} = \omega_a^{(i_a)} + \Delta\omega_a^{(i_a)}$ with equations (29)-(31);
    Update control vector $\boldsymbol{u}(k)$ through equations (27)-(28);
    Update $\hat{J}(k)$ through equations (20)-(22);
    Set $\Delta_a = \hat{J}(k)$, $i_a = i_a + 1$;
  end while
  Obtain the control vector $\boldsymbol{u}(k) = a(k)$ through equations (27)-(28);
  Update system state $\boldsymbol{x}(k+1) = f(\boldsymbol{x}(k), \boldsymbol{u}(k))$ of the EV, and calculate the unity function $U(\boldsymbol{x}(k), \boldsymbol{u}(k))$ through equation (35);
end for

V. SIMULATIONS

In this section, extensive simulations will be carried out to show the control performance of the proposed ADHDP-based Eco-ACC strategy. Firstly, the control performance of the proposed Eco-ACC strategy in terms of vehicle safety and tracking ability under UDDS, ARB02 and WLTC driving cycles is examined. Then, to show the robustness of the proposed Eco-ACC strategy in terms of extending battery life and improving energy economy, more simulations are carried out under more different typical driving cycles.

TABLE 1. Parameters of the learning process.

Parameter	Value
η_a: the learning rate of actor network	10^{-6}
η_c: the learning rate of critic network	10^{-6}
N_a: the number of neurons in the hidden laye of actor network	40
N_c: the number of neurons in the hidden layer of critic network	40
N_{am}: the maximum number of iterations in the actor network	50
N_{cm}: the maximum number of iterations in the critic network	50
T_a: the allowable error of actor network	10^{-4}
T_c: the allowable error of critic network	10^{-4}
γ: the discount factor of value function	1
α: the parameter of band-stop function	2
β: the parameter of band-stop function	8
$\hat{n}$: the paramer of band-stop function	1
cf: the compensating factor of band-stop function	2
α_1: the weight of unity function	1
α_2: the weight of unity function	1.2×10^{-6}
α_3: the weight of unity function	5×10^{-4}

TABLE 2. Main parameters of the EV model.

Parameter	Value	Unit
m	2530	kg
δ	1.08	–
η_t	0.98	–
R	0.345	m
μ	0.012	–
C_d	0.24	–
A_v	2.35	m^2
ρ	1.2	kg/m^3
G_r	9.73	–
g	9.81	N/kg

In the simulations, the parameters of the learning process are given in TABLE 1, and the main parameters of the EV in TABLE 2, which are used in [5], will also be employed here. And the initial states of the EV and its preceding vehicle are set as $s(0) = 0$ m, $s_p(0) = 10.5$ m, $v(0) = v_p(0) = 0$ m/s and $a_p(0) = 0$ m/s^2, and the initial control variable of the EV is obtained through policy iteration. It should be noted that, simulations are carried out in Matlab on a desktop computer with AMD Ryzen 5-3600 CPU and 16GB RAM, and the total iterative computation time of the critic and actor networks under UDDS, ARB02 and WLTC driving cycles is shown in TABLE 3, which shows the real-time performance of the proposed ADHDP-based Eco-ACC strategy.

When the preceding vehicle travels in the UDDS, ARB02 and WLTC driving cycles, the proposed Eco-ACC

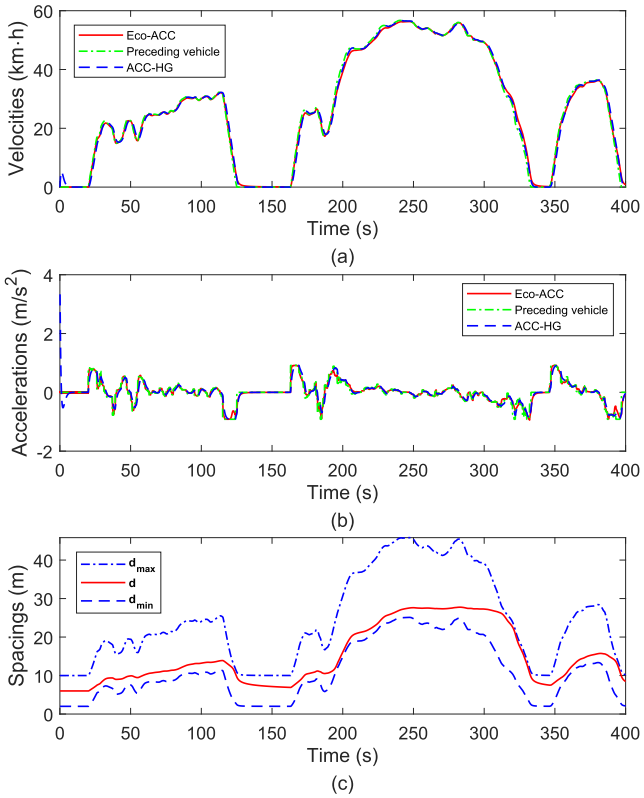


FIGURE 7. The evolutions of velocities, accelerations and spacings of the EV under UDDS driving cycle.

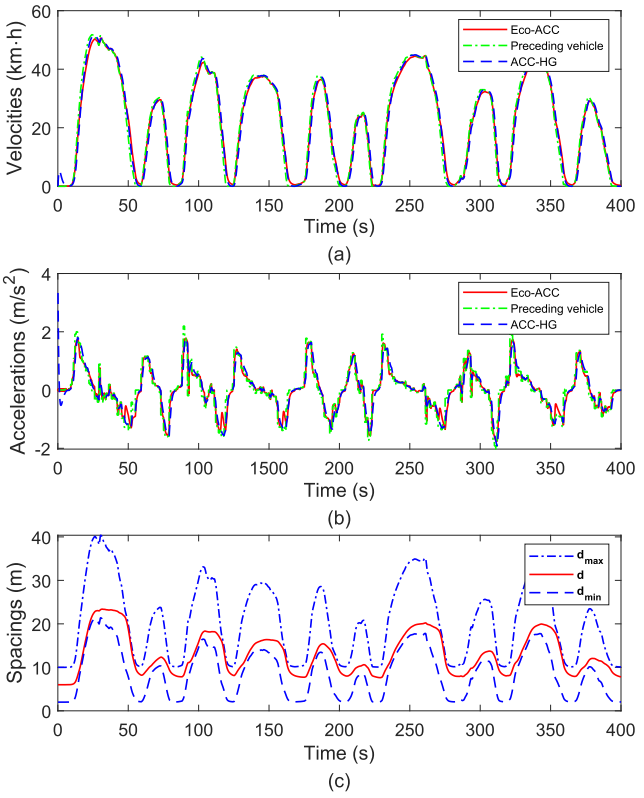


FIGURE 8. The evolutions of velocities, accelerations and spacings of the EV under ARB02 driving cycle.

TABLE 3. The iterative computation time of the critic and actor network under different driving cycles.

Driving cycles	Minimum computation time	Maximum computation time
UDDS	0.143 ms	11.462 ms
ARB02	0.103 ms	11.663 ms
WLTC	0.142 ms	10.586 ms

TABLE 4. Battery capacity losses (%) under different driving cycles.

Driving cycles	ACC-HG	Eco-ACC	Reduction
UDDS	9.26E-05	7.00E-05	24.41%
ARB02	7.43E-03	3.46E-03	53.43%
WLTC	2.14E-02	1.04E-02	51.40%
VAIL2NREL	5.83E-01	2.67E-01	54.20%
WVUSUB	3.92E-05	3.66E-05	6.63%
WVUCITY	1.68E-05	1.63E-05	2.98%

strategy and the ACC-HG strategy in [29] are used to control the EV, respectively. And the evolutions of velocity, acceleration, spacing of the EV under the corresponding driving cycles are obtained, which can be seen in FIGURE 7-9. It should be pointed out that in order to facilitate the observation of the simulation results, only the trajectories of the EV in the first 400 seconds are shown in the figures.

TABLE 5. Energy consumption (kw·h) under different driving cycles.

Driving cycles	ACC-HG	Eco-ACC	Reduction
UDDS	0.945	0.942	0.32%
ARB02	2.711	2.704	0.26%
WLTC	3.292	3.270	0.67%
VAIL2NREL	5.606	5.579	0.48%
WVUSUB	0.915	0.911	0.44%
WVUCITY	0.429	0.427	0.47%

As shown in FIGURE 7-9, the EV controlled by the proposed Eco-ACC can track the preceding vehicle well, and the speeds of the EV and its preceding vehicle are also the same; compared with ACC-HG, the acceleration of the EV controlled by Eco-ACC changes smoothly in the internal $[-2\ \text{m/s}^2, 2\ \text{m/s}^2]$, which can achieve riding comfort; when the EV is controlled by the Eco-ACC, the spacing between the EV and its preceding vehicle is always kept within the allowable range, which can ensure the vehicle safety.

In this paper, the dynamic battery capacity degradation model in [30] is used to measure the battery capacity loss.

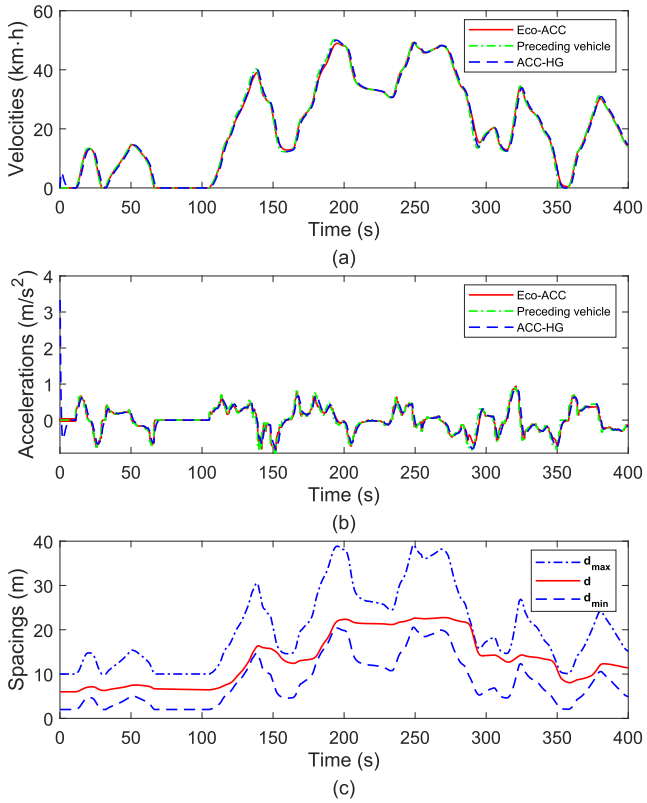


FIGURE 9. **The evolutions of velocities, accelerations and spacings of the EV under WLTC driving cycle.**

The battery capacity losses obtained by the proposed Eco-ACC and the ACC-HG method under various driving cycles, such as UDDS, ARB02, WLTC, can be seen in TBALE 4. Compared with the ACC-HG, the proposed Eco-ACC can decrease 2.98% $\sim$ 54.20% of the battery capacity loss, and thus significantly prolong the battery lifespan. It can be seen in TABLE 5 that, compared with the ACC-HG, the proposed Eco-ACC can decrease 0.26% $\sim$ 0.67% of the energy consumption, and thus slightly improve the energy economy.

VI. CONCLUSION

In order to achieve vehicle safety, extend battery lifespan, enhance riding comfort and improve energy economy, this paper proposes a model-free ADHDP-based Eco-ACC strategy for EVs in a car-following scenario. The simulation results under various driving cycles demonstrate that the presented Eco-ACC strategy can make the EV track its preceding vehicle well, significantly decrease the battery capacity loss, and slightly reduce the energy consumption. Furthermore, the proposed Eco-ACC is model-free and real-time, and can be robust in different car-following scenarios. With the rapid development of vehicle-to-vehicle communication technology, the cooperative ACC has attracted a lot of researchers due to its better control performance than the ACC [31]–[33]. In the future, the ADHDP-based method will be used to develop a novel CACC strategy for connected and automated vehicles to further improve the control performance.

REFERENCES

[1] Z. Xu, T. Wei, S. Easa, X. Zhao, and X. Qu, "Modeling relationship between truck fuel consumption and driving behavior using data from Internet of Vehicles," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 33, no. 3, pp. 209–219, Mar. 2018.

[2] C. Zhai, Y. Liu, and F. Luo, "A switched control strategy of heterogeneous vehicle platoon for multiple objectives with state constraints," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 5, pp. 1883–1896, May 2019.

[3] C. Zhai, X. Chen, C. Yan, Y. Liu, and H. Li, "Ecological cooperative adaptive cruise control for a heterogeneous platoon of heavy-duty vehicles with time delays," *IEEE Access*, vol. 8, pp. 146208–146219, 2020.

[4] C. Zhai, F. Luo, Y. Liu, and Z. Chen, "Ecological cooperative look-ahead control for automated vehicles travelling on freeways with varying slopes," *IEEE Trans. Veh. Technol.*, vol. 68, no. 2, pp. 1208–1221, Feb. 2019.

[5] C. Zhai, F. Luo, and Y. Liu, "Cooperative power split optimization for a group of intelligent electric vehicles travelling on a highway with varying slopes," *IEEE Trans. Intell. Transp. Syst.*, early access, Dec. 29, 2020, doi: 10.1109/TITS.2020.3045264.

[6] K. C. Dey, L. Yan, X. Wang, Y. Wang, H. Shen, M. Chowdhury, L. Yu, C. Qiu, and V. Soundararaj, "A review of communication, driver characteristics, and controls aspects of cooperative adaptive cruise control (CACC)," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 2, pp. 491–509, Feb. 2016.

[7] J. Zhou and H. Peng, "Range policy of adaptive cruise control vehicles for improved flow stability and string stability," *IEEE Trans. Intell. Transp. Syst.*, vol. 6, no. 2, pp. 229–237, Jun. 2005.

[8] G. Marsden, M. McDonald, and M. Brackstone, "Towards an understanding of adaptive cruise control," *Transp. Res. C, Emerg. Technol.*, vol. 9, no. 1, pp. 33–51, Feb. 2001.

[9] B. Sakhdari and N. L. Azad, "Adaptive tube-based nonlinear MPC for economic autonomous cruise control of plug-in hybrid electric vehicles," *IEEE Trans. Veh. Technol.*, vol. 67, no. 12, pp. 11390–11401, Dec. 2018.

[10] M. Wang, H. Yu, G. Dong, and M. Huang, "Dual-mode adaptive cruise control strategy based on model predictive control and neural network for pure electric vehicles," in *Proc. 5th Int. Conf. Transp. Inf. Saf. (ICTIS)*, Jul. 2019, pp. 1220–1225.

[11] S. Zhang and X. Zhuan, "Multi-objective optimization for pure electric vehicle during a car-following process," in *Proc. Chin. Control Conf. (CCC)*, Jul. 2019, pp. 2884–2888.

[12] J. Guo, W. Li, J. Wang, and K. Li, "Multi-objective cooperative control of intelligent electric vehicles with adaptive cruise and regenerative braking," *Automot. Eng.*, vol. 42, no. 12, pp. 1638–1646, Dec. 2020.

[13] B. Bayar, S. A. Sajadi-Alamdari, F. Viti, and H. Voos, "Impact of different spacing policies for adaptive cruise control on traffic and energy consumption of electric vehicles," in *Proc. 24th Medit. Conf. Control Autom. (MED)*, Jun. 2016, pp. 1349–1354.

[14] J. N. Barkenbus, "Eco-driving: An overlooked climate change initiative," *Energy Policy*, vol. 38, no. 2, pp. 762–769, Feb. 2010.

[15] H. Liu, C. Miao, and G. G. Zhu, "Economic adaptive cruise control for a power split hybrid electric vehicle," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 10, pp. 4161–4170, Oct. 2020.

[16] S. E. Li, Q. Guo, L. Xin, B. Cheng, and K. Li, "Fuel-saving servo-loop control for an adaptive cruise control system of road vehicles with step-gear transmission," *IEEE Trans. Veh. Technol.*, vol. 66, no. 3, pp. 2033–2043, Mar. 2017.

[17] Y. Jia, R. Jibrin, and D. Görges, "Energy-optimal adaptive cruise control for electric vehicles based on linear and nonlinear model predictive control," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 14173–14187, Dec. 2020.

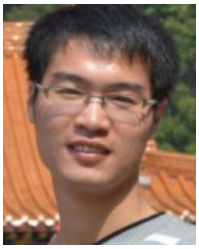
[18] J. Borek, B. Groelke, C. Earnhardt, and C. Vermillion, "Economic optimal control for minimizing fuel consumption of heavy-duty trucks in a highway environment," *IEEE Trans. Control Syst. Technol.*, vol. 28, no. 5, pp. 1652–1664, Sep. 2020.

[19] D. Moser, R. Schmied, H. Waschl, and L. del Re, "Flexible spacing adaptive cruise control using stochastic model predictive control," *IEEE Trans. Control Syst. Technol.*, vol. 26, no. 1, pp. 114–127, Jan. 2018.

[20] H. Lee, N. Kim, and S. W. Cha, "Model-based reinforcement learning for eco-driving control of electric vehicles," *IEEE Access*, vol. 8, pp. 202886–202896, 2020.
[21] G. Li and D. Görges, "Ecological adaptive cruise control for vehicles with step-gear transmission based on reinforcement learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 11, pp. 4895–4905, Nov. 2020.
[22] R. Raustad, "Electric vehicle life cycle cost analysis," Electr. Vehicle Transp. Center, Tallahassee, FL, USA, Tech. Rep. FSEC-CR-2053-17, 2017.
[23] T. A. Burress, S. L. Campbell, C. L. Coomer, C. W. Ayers, A. A. Wereszczak, J. P. Cunningham, L. D. Marlino, L. E. Seiber, and H. T. Lin, "Evaluation of the 2010 Toyota Prius hybrid synergy drive system," Oak Ridge Nat. Lab., Oak Ridge, TN, USA, Tech. Rep. ORNL/TM-2010/253, 2011.
[24] L. Li, X. Wang, and J. Song, "Fuel consumption optimization for smart hybrid electric vehicle during a car-following process," *Mech. Syst. Signal Process.*, vol. 87, pp. 17–29, Mar. 2017.
[25] P. Golchoubian and N. L. Azad, "Real-time nonlinear model predictive control of a battery–supercapacitor hybrid energy storage system in electric vehicles," *IEEE Trans. Veh. Technol.*, vol. 66, no. 11, pp. 9678–9688, Nov. 2017.
[26] J. Si and Y.-T. Wang, "Online learning control by association and reinforcement," *IEEE Trans. Neural Netw.*, vol. 12, no. 2, pp. 264–276, Mar. 2001.
[27] F. L. Lewis, D. Vrabie, and K. G. Vamvoudakis, "Reinforcement learning and feedback control: Using natural decision methods to design optimal adaptive controllers," *IEEE Control Syst.*, vol. 32, no. 6, pp. 76–105, Dec. 2012.
[28] C. Zhai, F. Luo, and Y. Liu, "A novel predictive energy management strategy for electric vehicles based on velocity prediction," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 12559–12569, Nov. 2020.
[29] C. Zhai, F. Luo, and Y. Liu, "Cooperative look-ahead control of vehicle platoon for maximizing fuel efficiency under system constraints," *IEEE Access*, vol. 6, pp. 37700–37714, 2018.
[30] Z. Song, H. Hofmann, J. Li, X. Han, and M. Ouyang, "Optimization for a hybrid energy storage system in electric vehicles using dynamic programing approach," *Appl. Energy*, vol. 139, pp. 151–162, Feb. 2015.
[31] J. C. Zegers, E. Semsar-Kazerooni, J. Ploeg, N. van de Wouw, and H. Nijmeijer, "Consensus control for vehicular platooning with velocity constraints," *IEEE Trans. Control Syst. Technol.*, vol. 26, no. 5, pp. 1592–1605, Sep. 2018.
[32] S. Baldi, D. Liu, V. Jain, and W. Yu, "Establishing platoons of bidirectional cooperative vehicles with engine limits and uncertain dynamics," *IEEE Trans. Intell. Transp. Syst.*, vol. 42, no. 5, pp. 679–2691, Feb. 2021.
[33] H. Xing, J. Ploeg, and H. Nijmeijer, "Compensation of communication delays in a cooperative ACC system," *IEEE Trans. Veh. Technol.*, vol. 69, no. 2, pp. 1177–1189, Feb. 2020.

XIYAN CHEN was born in Wenzhou, Zhejiang, China. She is currently pursuing the B.S. degree with Hangzhou Dianzi University, Hangzhou, China. Her main research interests include reinforcement learning, autonomous vehicle control, and cooperative control.

JIAN YANG was born in Guangan, Sichuan, China. He received the B.S. degree in computer science and information engineering from Chongqing Technology and Business University, China, in 2019. He is currently pursuing the M.S. degree with Hangzhou Dianzi University, Hangzhou, China. He main research interests include reinforcement learning and intelligent driving.



CHUNJIE ZHAI (Member, IEEE) was born in Bozhou, Anhui, China. He received the B.S. degree from the Jiangxi University of Science and Technology, Ganzhou, China, in 2012, and the Ph.D. degree from the South China University of Technology, Guangzhou, China, in 2019. He is currently a Lecturer with the School of Automation, Hangzhou Dianzi University, Hangzhou, China. He has published more than 15 SCI/EI academic articles in international journals and conferences. His main research interests include autonomous vehicle control, cooperative control, and intelligent transportation systems.

JIEDONG LOU was born in Shaoxing, Zhejiang, China. He received the B.S. degree from Hangzhou Dianzi University, Hangzhou, China, in 2019, where he is currently pursuing the master's degree. His main research interests include autonomous vehicle control, reinforcement learning, and cooperative control.

CHENGGANG YAN received the B.S. degree in computer science from Shandong University, Jinan, China, in 2008, and the Ph.D. degree in computer science from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2013. He is currently a Professor with Hangzhou Dianzi University, Hangzhou, China. Before that, he was an Assistant Research Fellow with Tsinghua University. He has authored or coauthored more than 30 refereed journal and conference articles. His research interests include machine learning, image processing, computational biology, and computational photography. As the coauthor, he received the Best Paper Awards in International Conference on Game Theory for Networks 2014, the SPIE/COS Photonics Asia Conference 9273 2014, and the Best Paper Candidate in International Conference on Multimedia and Expo 2011.