

## Article

# Hierarchical Optimization Based on Deep Reinforcement Learning for Connected Fuel Cell Hybrid Vehicles through Signalized Intersections

Hongquan Dong \*, Lingying Zhao, Hao Zhou and Haolin Li

School of Mechanical Engineering, Taiyuan University of Science and Technology, Taiyuan 030024, China; zhaoly2312@163.com (L.Z.); 2020092@tyust.edu.cn (H.Z.); lhl2286@163.com (H.L.)

\* Correspondence: donghongquan@tyust.edu.cn; Tel.: +86-351-699-8115

**Abstract:** With the advantages of non-pollution and energy-saving, hydrogen fuel cell hybrid vehicles (HFCHVs) are regarded as one of the potential traveling ways in the future. The energy management of FCHVs has a huge energy-efficient potential which is combined with the Internet of Things (IOT) and auto-driving technologies. In this paper, a hierarchical joint optimization method that combines deep deterministic policy gradient and dynamic programming (DDPG-DP) for speed planning and energy management of the HFCHV is proposed for urban road driving scenarios. The results demonstrate that when the HFCHV is operating in driving scenario 1, the traveling efficiency of the DDPG-DP algorithm is 17.8% higher than that of the IDM-DP algorithm, and the hydrogen fuel consumption is reduced by 2.7%. In contrast, the difference in the traveling efficiency and fuel economy is small among the three algorithms in driving scenario 2, the number of idling/stop situations of the DDPG-DP algorithm is reduced compared with that of the IDM-DP algorithm. This will support further research for multi-objective eco-driving optimization of fuel cell hybrid vehicles.

**Keywords:** hydrogen fuel cell hybrid vehicle; hierarchical joint optimization; traveling efficiency; fuel economy; urban roads driving scenario



**Citation:** Dong, H.; Zhao, L.; Zhou, H.; Li, H. Hierarchical Optimization Based on Deep Reinforcement Learning for Connected Fuel Cell Hybrid Vehicles through Signalized Intersections. *Processes* **2023**, *11*, 2689. <https://doi.org/10.3390/pr11092689>

Academic Editor: Takuya Oda

Received: 13 August 2023

Revised: 30 August 2023

Accepted: 4 September 2023

Published: 7 September 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

With the advent of connected vehicle technologies, connected vehicles (CVs) have a great potential to significantly increase road safety [1], ride comfort, traffic efficiency [2,3], and energy efficiency. At the same time, combined with energy-saving optimization of the vehicle driving processes, green and safe travel has been achieved [4,5].

Hydrogen fuel cells have emerged as a crucial component for the future development of hybrid vehicles with high energy efficiency due to their high efficiency, energy savings and non-polluting advantages [6,7]. In order to enhance the vehicle's driving requirements and reduce operating costs, hydrogen fuel cell hybrid vehicles (HFCHVs) currently on the market basically use a powertrain of hybrid fuel cells and lithium batteries [8]. Compared with the single power system using hydrogen fuel cells in the vehicle, the hybrid power system has a more complex structure and the power source energy management strategy needs to be further optimized [9].

### 1.1. Energy Management of Literature Review

Currently, rule-based, optimization-based and learning-based strategies are the three main types of energy management strategies used in hybrid vehicles. Rule-based energy management strategies include fuzzy rule control, which is simple and practical, but often not optimal [10]. Optimization-based energy management strategies are divided into global optimization and transient optimization. Then, as representative algorithms for global optimization, dynamic programming algorithms mainly achieve global optimization based on global driving information [11]. Transient optimization algorithms are mainly

represented by equivalent fuel consumption minimization (ECMS) algorithms, which can solve the optimal results from local driving information [12,13]. In recent years, due to the self-learning and adaptation of learning algorithms for intelligent agents, learning-based energy management strategies have been widely studied, which include deep reinforcement learning algorithms and neural network learning algorithms [14,15]. Chen et al. [16] constructed an online intelligent energy management controller based on two neural network (NN) modules to improve the fuel economy of vehicles. Jia et al. [17] proposed a novel energy management strategy for hybrid electric buses with fuel cell health and battery thermal- and health-constrained awareness to decrease battery aging and total operating costs. Tang et al. [18] proposed an energy management strategy for fuel cell hybrid vehicles based on the deep Q-learning algorithm, where the lifetime of the fuel cell and lithium battery was taken as an important objective during the optimizing process [19]. Zhang et al. [20] presented a new power management strategy to deal with the inaccuracies and uncertainties of terrain information and extend the battery life. Yang et al. [21] developed a new method for calculating the fuel conversion factor of hybrid vehicles to maintain the battery state of charge fluctuations within a reasonable range. In the urban road driving scenarios, Jia et al. [22] applied a health-aware energy management strategy for fuel cell hybrid buses considering air-conditioning control based on the TD3 algorithm to decrease the driving economy costs. Anselma et al. [23] proposed an energy management strategy for hybrid vehicles based on causal optimal control to reduce fuel consumption while maintaining the battery state of charge.

### *1.2. Speed Planning of Literature Review*

Currently, eco-driving of connected electric vehicles and connected fuel vehicles usually only optimizes vehicle speed planning, and energy consumption assessment is usually based on engine fuel consumption models or battery consumption models [24–27]. However, due to the fact that the powertrains of hybrid vehicles usually consist of different power sources, the eco-driving problem for hybrid vehicles requires coordinated optimization of speed planning and energy management [28–30]. In terms of research on eco-driving for hybrid vehicles, Bai et al. [31] proposed a hybrid reinforcement learning framework to enable eco-driving at signalized intersections in mixed traffic scenarios. Wang et al. [32] introduced a data-driven predictive control strategy for optimal control of connected vehicles in mixed traffic scenarios. Liu et al. [33] presented a bilaterally convex method to achieve eco-driving of fuel cell hybrid vehicles through signalized intersection scenarios. Dong et al. [34] proposed a predictive energy-efficient driving strategy for connected electric vehicles in signalized intersection scenarios to reduce energy consumption and battery life loss.

### *1.3. Research Motivation and Contribution*

The above-mentioned literature revealed that the research of connected hybrid vehicles needs to be deepened, and the multi-objective optimization of connected vehicles for traveling efficiency and energy-saving driving has significant research value. Therefore, this paper carries out a hierarchical multi-objective optimization research on the eco-driving of speed planning and energy management for HFCHVs in urban road driving scenarios. This study will have extremely important implications for the energy economy cost and emission reduction of vehicles and traffic optimization. The relevant contributions of this paper are specified below:

- (1) In order to optimize the speed of connected vehicles, this paper proposes a DDPG algorithm that improves the traveling efficiency of the HFCHV through traffic lights. When passing through multiple traffic lights smoothly, the HFCHV using the DDPG algorithm greatly reduces the traveling time in driving scenario 1, compared with the IDM algorithm. The HFCHV using the DDPG algorithm has only a 5 s difference in traveling time compared with the DP algorithm. Therefore, the DDPG algorithm proposed in this paper enables the HFCHV to have excellent traveling efficiency.

(2) In this paper, a multi-objective hierarchical optimization framework that combines deep deterministic policy gradient and dynamic programming (DDPG-DP) for speed planning and energy management of the HFCHV is proposed. During the driving scenario 1, the fuel economy of the HFCHV using the DDPG-DP algorithm is 2.7% higher than that of the IDM-DP algorithm, which can achieve the multi-objective optimization of traveling efficiency and energy conservation for the hydrogen fuel cell hybrid vehicle.

The structure of this paper is as follows. In Section 2, the modeling of the powertrain of the HFCHV and the driving scenario simulation models with traffic lights are described in detail. In Section 3, the eco-driving problem for HFCHV is presented along with a framework for a collaborative optimization strategy. The upper-level speed planning and the lower-level energy management are discussed and described in detail in Sections 4 and 5, respectively. The eco-driving findings are shown in Section 6, and the conclusion is shown in Section 7.

## 2. Vehicles Modeling and Signal Timing

The power system of the HFCHV is usually composed of a lithium battery and hydrogen fuel cell, and its structure topology is shown in Figure 1. The motor that powers the vehicle through the transmission system is powered by the lithium-ion battery and the fuel cell, which are combined and connected to the vehicle by a DC/DC converter. The detailed parameters of HFCHV are shown in Table 1.

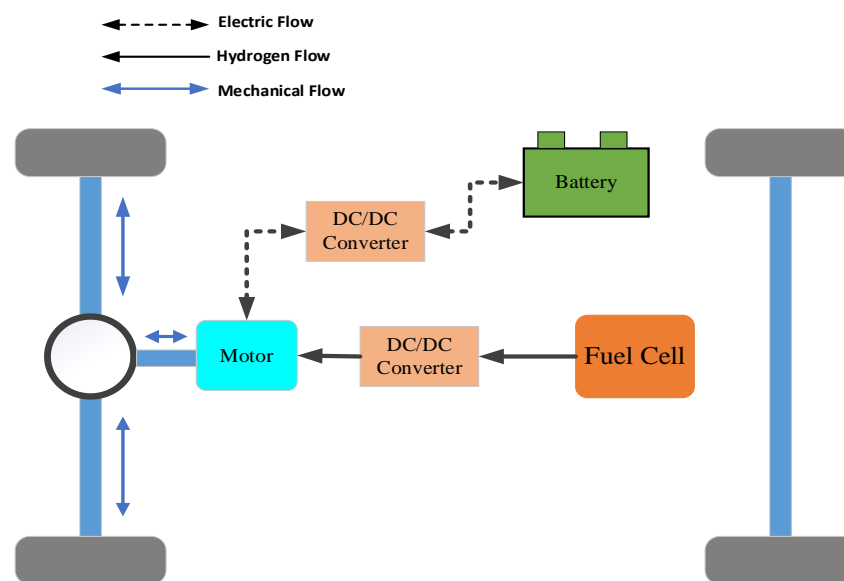


Figure 1. Power system topology of HFCHV.

Table 1. Vehicle parameters of HFCHV.

| Items               | Parameters (Unit)                            | Value |
|---------------------|--|-------|
| Vehicle             | Crub weight $M$ (kg)                         | 1380  |
|                     | Rolling resistance coefficient $f$           | 0.013 |
|                     | Wheel radius (m)                             | 0.282 |
|                     | Air mass density $\rho$ (kg/m <sup>3</sup> ) | 1.2   |
|                     | Vehicle frontal area $A_r$ (m <sup>2</sup> ) | 2.23  |
| Fuel cell           | Maximum efficiency                           | 0.596 |
|                     | Maximum power (kW)                           | 50    |
| Lithium-ion battery | Capacity (Ah)                                | 26    |
|                     | Maximum power (kW)                           | 85    |
| Motor               | Maximum efficiency                           | 0.92  |
|                     | Maximum torque (N·m)                         | 271   |

The longitudinal dynamics of the vehicle are as follows:

$$P_{req} = [m \cdot \left(\frac{dv}{dt}\right) + 0.5\rho A_r c_d v^2 + mg \sin \theta + mgf \cos \theta]v \quad (1)$$

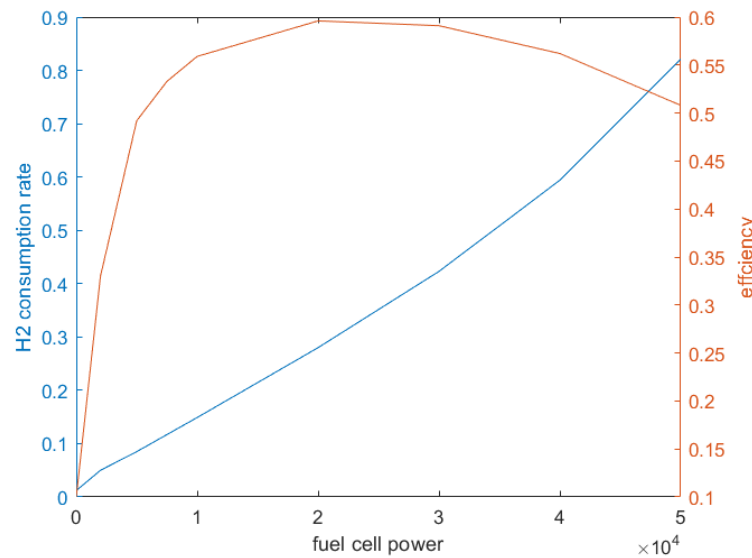
where,  $P_{req}$  is the power demanded by the vehicle,  $m$  is the mass of the HFCHV, and  $\rho$ ,  $A_r$ ,  $c_d$  denotes the air density, windward area, and air resistance coefficient, respectively. This paper mainly studies in a flat and straight single-road environment, so  $\theta$  is set to zero.

### 2.1. Hydrogen Fuel Cell

Because hydrogen fuel cell is a complex nonlinear system, and empirical data are commonly used in engineering models to fit the output characteristics of hydrogen fuel cells. The energy conversion process of the hydrogen fuel cell system is described by the following equation:

$$P_{fc} = \dot{m}_{H_2} LHV_{H_2} \eta_{fc} \quad (2)$$

where,  $\eta_{fc}$  is the efficiency of the hydrogen fuel cell hybrid power system,  $\dot{m}_{H_2}$  is the hydrogen consumption rate and  $LHV_{H_2}$  is the low calorific value of hydrogen, the relationship between instantaneous hydrogen consumption and power for this experiment is shown in Figure 2.



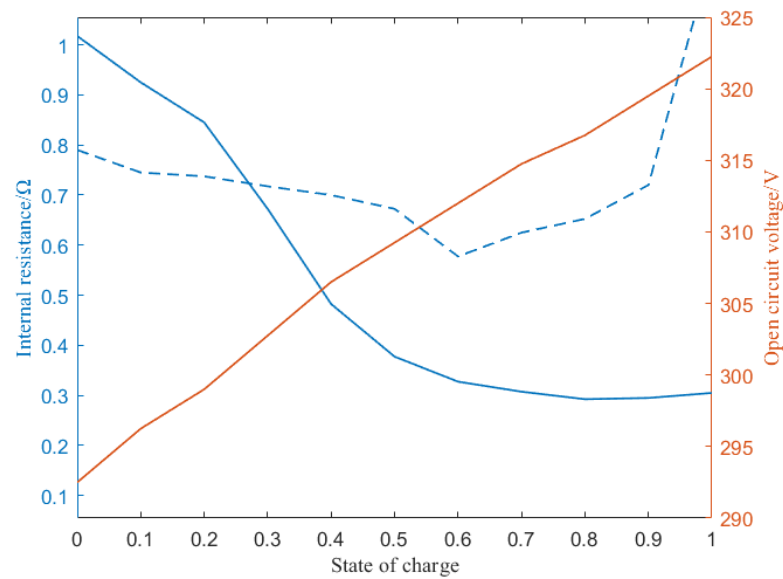
**Figure 2.** Hydrogen consumption rate and efficiency for hydrogen fuel cell. (The blue line represents hydrogen consumption rate, the orange line represents the efficiency of hydrogen fuel cell).

### 2.2. Lithium Battery

In the hydrogen fuel cell hybrid power system, the lithium-ion battery can not only cooperate with the hydrogen fuel cell to produce electrical energy to meet the power demand of the whole vehicle but also be able to recycle the energy generated by the braking of the HFCHV during the deceleration process to reduce the loss of energy consumption. The equivalent circuit model of the lithium battery is as follows:

$$\begin{cases} P_{bat} = P_{req} - P_{fc} \\ I_{bat} = \frac{V - \sqrt{V^2 - 4R_{Int}P_{bat}}}{2R_{Int}} \\ d(SOC) = -\frac{I_{bat}}{Q_0} dt \end{cases} \quad (3)$$

where,  $V$  and  $R_{Int}$  are the open-circuit voltage and internal resistance of the power cell, respectively. The relationship between the internal resistance of charge-discharge, open circuit voltage and state of charge (SOC) of the lithium battery model is shown in Figure 3.



**Figure 3.** Relationship between internal resistance of charge-discharge, open-circuit voltage and SOC of lithium battery. (The blue solid line represents internal resistance of charge-discharge, the orange solid line represents the open-circuit voltage, and the dashed line represents the SOC of lithium battery).

### 2.3. Modeling of Traffic Light Signal

In this section, a fixed signal disturbing model is developed. There are  $n$  signals on a straight road of length  $L$  and the distance from the  $i$ th signal to the starting point is  $L_i$ ,  $L_i \in [0, L]$ , and  $i = \{1, 2, 3, \dots, n\}$ . The cycle duration of the  $i$ th signal is defined as  $T^i$ .

$$T_R^i + T_G^i = T^i \quad (4)$$

When the vehicle arrives at the  $i$ th signal from the starting point, the time during which the  $i$ th signal operates in its own signal cycle  $T_d^i(t)$  can be defined as:

$$T_d^i(t) = (T_0^i + t) \bmod T^i \quad (5)$$

where,  $T_R^i$ ,  $T_G^i$  denote the duration of the  $i$ th red light and green light, respectively,  $T_0^i$  denotes the initial operating time of the  $i$ th traffic light when the vehicle departs, and  $t$  denotes the absolute driving time of the vehicle.

In order to avoid red-light running by the vehicle, in the next subsequent algorithm, the driving time constraints of the vehicle are incorporated as follows:

$$T_d^i(t_d^i) \geq T_R^i \quad (6)$$

where,  $t_d^i$  is the absolute driving time of the vehicle through the  $i$ th traffic light.

### 3. Analysis of Eco-Driving Problems

Eco-driving optimization for the HFCHV includes upper-level vehicle speed planning and lower-level energy management control of energy storage systems [33]. The goal of speed planning is to maximize travel efficiency and minimize the energy consumption of the vehicle passing through all traffic lights. Based on the upper-level optimized results, the lower-level energy management strategy optimizes the algorithm on the obtained vehicle driving data, in order to obtain the optimal outputs from the various power systems. Combined with the analysis results of the speed planning and energy management, the

constraints of the relevant algorithms designed subsequently in this paper are specifically shown in Equations (7) and (8).

$$\begin{cases} v_k = v_{k-1} + a \\ s_k = s_{k-1} + v_k \Delta t \\ t(0) = 0, t(s_k) = t_k \\ v(0) = v_0, v(s_k) = v_k \\ v_{\min} \leq v(s_k) \leq v_{\max} \\ t_l^i(t_k^i) \geq T_r^i \end{cases} \quad (7)$$

$$\begin{cases} SOC(0) = SOC_0, SOC(end) = SOC_0 \\ SOC_{\min} \leq SOC(t) \leq SOC_{\max} \\ P_{fc_{\min}} \leq P_{fc}(t) \leq P_{fc_{\max}} \\ P_{bat_{\min}} \leq P_{bat}(t) \leq P_{bat_{\max}} \end{cases} \quad (8)$$

where,  $v_k, v_{k-1}, s_k, s_{k-1}$  represents the vehicle speed and driving distance in the  $k$ th,  $(k-1)$ th time,  $v_0$  is the vehicle speed at the starting timing, which is set to 0 in this paper. In order to comply with the traffic rules in the urban road, the maximum and minimum speeds of the vehicle at the position  $s_k$  are set to 60 km/h and 0 km/h, respectively.  $SOC_{\min}, SOC_{\max}, P_{fc_{\min}}, P_{fc_{\max}}, P_{bat_{\min}}$  and  $P_{bat_{\max}}$  represents 0.3, 0.8, 0 kW, 50 kW,  $-85$  kW and 85 kW, respectively.

#### 4. Upper-Level Speed Planning Algorithm

The upper-level speed planning problem for the HFCHV passing through traffic light intersection scenarios includes vehicle dynamics, signal lights and traffic condition constraints. In this section, the DDPG algorithm is proposed to optimize the energy-saving potential and traveling efficiency of the HFCHV passing through traffic light intersection scenarios, and the Intelligent Driver Model (IDM) algorithm and the DP (Dynamic Programming) algorithm are introduced, respectively.

##### 4.1. IDM Algorithm

The IDM plans the acceleration of the vehicle by acquiring information about the vehicle and the traffic lights to ensure that the vehicle passes through the traffic lights safely. As one of the classical algorithms for speed planning, the acceleration control strategy of the IDM model is as follows:

$$a_{IDM} = a_{\max} \left[ 1 - \left( \frac{v(t)}{v_{\max}} \right)^4 - \left( \frac{L^*(v, \Delta v)}{\Delta x} \right)^2 \right] \quad (9)$$

where,  $\Delta v$  and  $\Delta x$  are the speed difference and the relative distance to the previous vehicle, respectively, and  $L^*(v, \Delta v)$  is the desired distance, which is defined as follows:

$$L^*(v, \Delta v) = l_0 + \max \left[ 0, vt_x + \frac{v \cdot \Delta v}{2\sqrt{a_{\max} \cdot b}} \right] \quad (10)$$

where,  $l_0$  is the minimum vehicle spacing distance, which is set to 2 m to ensure safe driving,  $t_x$  is the desired time interval, which is set to 1 s, and  $b$  is the desired deceleration rate.

In order to safely pass through the upcoming traffic light intersection, this paper considers the traffic light ahead as the front vehicle and assumes that the vehicle can preview the traffic signal state within the signal sensing range. The acceleration of the modified IDM model is reformulated as follows:

$$a_{I-IDM} = \begin{cases} a_{IDM}, P_{iss} = 0 \\ -\frac{v_k^2}{2\Delta x}, P_{iss} = 1 \end{cases} \quad (11)$$

where  $P_{iss}$  is the traffic signal status, 1 means red light, 0 means green light or that the vehicle is outside the signal sensing range.

#### 4.2. DDPG Algorithm

Based on the nonlinear fitting ability and self-learning ability of deep reinforcement learning, deep reinforcement learning has been widely researched on the traveling problems at signalized intersection scenarios. Deep reinforcement learning algorithms are typically employed to improve vehicle fuel economy, driving safety and traffic efficiency, according to the signal phase and timing information of traffic lights.

On the basis of deep Q-learning, the DDPG algorithm applies the Actor network  $\mu$  to fit the policy function to output continuous action. The Actor network  $\mu$  interacts with the environment to generate action  $a_t$ , Critic network  $Q$  evaluates the performance of the Actor network combined with Bellman's equation and generates  $a_{t+1}$ . The DDPG algorithm adopts the Actor-Critic network framework to update the Actor network and Critic network parameters, which is shown in Equation (12).

$$\begin{cases} \theta^{Q'} \leftarrow \tau\theta^Q + (1 - \tau)\theta^{Q'} \\ \theta^{\mu'} \leftarrow \tau\theta^\mu + (1 - \tau)\theta^{\mu'} \end{cases} \quad (12)$$

where  $\theta^\mu$  represents the parameters of the Actor network,  $\theta^Q$  represents the parameters of the Critic network.

The Critic network is updated by minimizing the Loss function, the Actor network selects the optimal action policy by gradient update. The gradient of objective function  $J(\theta^\mu)$  and Loss function  $L(\theta^Q)$  are described in the following Equation (13):

$$\begin{cases} L(\theta^Q) = \frac{1}{N} \sum_t \left( r_t + \gamma Q'(s_{t+1}, \mu'(s_{t+1} | \theta^{\mu'})) | \theta^{Q'} - Q(s_t, a_t | \theta^Q) \right)^2 \\ \nabla_{\theta^\mu} J \approx E \left[ \nabla_{\theta^\mu} Q(s, a | \theta^Q) | s = s_t, a = u(s_t | \theta^\mu) \right] = \\ E \left[ \nabla_{\theta^\mu} Q(s, a | \theta^Q) | s = s_t, a = u(s_t) \nabla_{\theta^\mu} \mu(s | \theta^\mu) |_{s=s_t} \right] \end{cases} \quad (13)$$

The pseudocode of the DDPG Algorithm 1 is shown as follows:

---

#### Algorithm 1 DDPG

---

- 1: Initialization : critic network and actor network with weights  $\theta^Q$  and  $\theta^\mu$ , target network  $Q'$  and  $\mu'$  with weights  $\theta^{Q'} \leftarrow \theta^Q$ ,  $\theta^{\mu'} \leftarrow \theta^\mu$ , memory pool R
  - 2: For episode: 1:M do
  - 3: Get initial states:  $s_0$
  - 4: For  $t < s_{\max}$  do
  - 5: Choose action  $a_t = \mu(s_t | \theta^\mu)$  according to the policy and state
  - 6: Choose action  $a_t = \mu(s_t | \theta^\mu)$  according to the policy and state
  - 7: Execute action  $a_t$ , reward  $r_t$  and next state  $s_{t+1}$
  - 8: Store transition  $(s_t, a_t, r_t, s_{t+1})$  in memory pool
  - 9: Store transition  $(s_t, a_t, r_t, s_{t+1})$  in memory pool
  - 10: Sampling a mini\_batch of transition from memory pool R
  - 11: Set  $y_t = r_t + \gamma Q'(s_{t+1}, \mu'(s_{t+1} | \theta^{\mu'})) | \theta^{Q'}$
  - 12: Updating the critic network parameters with minimizing the Loss:  
 $L(\theta^Q) = \frac{1}{N} \sum_t \left( r_t + \gamma Q'(s_{t+1}, \mu'(s_{t+1} | \theta^{\mu'})) | \theta^{Q'} - Q(s_t, a_t | \theta^Q) \right)^2$
  - 13: Updating the actor network parameters :  $\begin{cases} \theta^{Q'} \leftarrow \tau\theta^Q + (1 - \tau)\theta^{Q'} \\ \theta^{\mu'} \leftarrow \tau\theta^\mu + (1 - \tau)\theta^{\mu'} \end{cases}$
  - 14: if vehicle passed all traffic lights successfully:  $t = e10$
  - 15: End
  - 16: End
- 

In a single-lane training environment with many traffic light intersection scenarios, the DDPG intelligent agents take action based on the vehicle state information which is

observed at each discrete time step  $\Delta t$ . In order to satisfy the real road driving requirements, the speed limit of the vehicle is 0 m/s to 60 km/h. Meanwhile, in order to ensure the driving comfort of the driver, the acceleration limit of the vehicle is  $-2.0 \text{ m/s}^2$  to  $1.4 \text{ m/s}^2$ . As for the speed planning problem through traffic lights, the state space of the DDPG algorithm is set to  $X = \{v \ s\}$ ,  $s$  is the driving distance of the vehicle from the starting point,  $s_{max}$  is the maximum length of the road to be traveled, and the termination condition of the driving process through traffic lights is  $s \geq s_{max}$ , and the action space is set to the acceleration  $a$  of the vehicle. At the same time, a reward function for multi-objective optimization is also designed, and the formula is as follows:

$$\begin{cases} R = \omega_1 \cdot r_1 + \omega_2 \cdot r_2 + r_3 \\ r_1 = |\dot{m}_{fc}| \\ r_2 = \text{abs}(v(t) - v_{\max})/5 \\ r_3 = \begin{cases} -100, P(t_p) = 0 \\ 100, P(t_p) = 1 \end{cases} \end{cases} \quad (14)$$

where  $r_1$  guides the agent to improve energy saving when passing multiple traffic lights,  $r_2$  mainly guides the vehicle to be able to pass the traffic lights quickly,  $r_3$  guides the agent to control the vehicle to pass the traffic lights smoothly without violating the traffic rules, and  $t_p$  is the time for the vehicle to pass the traffic lights.

#### 4.3. DP Algorithm

In this section, as a classic global algorithm, the dynamic programming algorithm is used as the optimal benchmark algorithm, and the multi-objective function of the dynamic planning algorithm is set as follows:

$$J = \int_0^{t_M} M(av + mgfv + 0.5 \frac{\rho A_r c_d}{M} v^3) + \varepsilon |v - v_{\max}| dt \quad (15)$$

The state equation of the dynamic programming algorithm is as follows:

$$\begin{cases} v(l+1) = v(l) + a(l) \\ t(l+1) = t(l) + \frac{1}{v(l+1)} \end{cases} \quad (16)$$

where, with the condition of ensuring the calculation accuracy of the algorithm, the simulation step of the algorithm  $l$  is set to 1 m.

### 5. Energy Management

In order to obtain the online driving data information by the upper-level speed planning algorithm, this section uses the Dynamic Programming algorithm to solve the energy management problem of the HFCHV, in order to obtain the optimal energy-saving results. Based on the constraints related to the energy management problem of the HFCHV in Section 3, the state variable of the DP algorithm is the SOC of the lithium battery, and the action variable is the output power of the hydrogen fuel cell, and the objective function of the hydrogen consumption is established as follows.

$$J = \min \int_0^{t_M} \dot{m}_{fc} dt \quad (17)$$

In this section, the energy management control strategy of the DP algorithm is used for the maintenance simulation of lithium battery power. In the DP algorithm, the initial and final SOC values of the power cell are set to 0.6 to minimize the hydrogen consumption of the hydrogen fuel cell during the vehicle operation process.



## 6. Results and Discussion

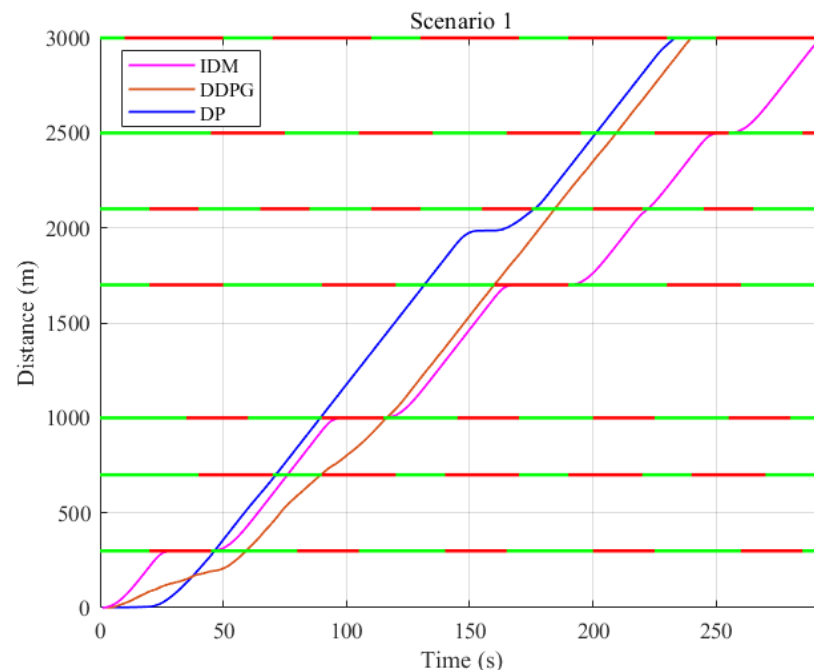
### 6.1. Test Scenarios and Trajectories

In this section, eco-driving simulation tests are conducted on three hierarchical optimization algorithms based on the MATLAB (2019 version) and pycharm simulation platforms. The global optimality of the DP algorithm is used as a benchmarking method for both the upper-level speed planning and the hierarchical optimization algorithms. In order to facilitate the subsequent analysis and study of the vehicle during the actual operation process, this study is divided into two scenarios, where the parameter settings of driving scenario 1 and driving scenario 2 are shown in Table 2, respectively. To verify the performance of the proposed algorithm in different driving scenarios, two different urban roads are chosen in this section. The position of the vehicle is assumed as 0 m in the two driving scenarios, the position  $S$  represents the distance of the  $i$ th traffic light. The green light duration is  $T_G$ , the red light duration is  $T_R$ , the initial periodic time  $t_0$  represents the red signal period, the green light period and the passed green light period at the start of the  $i$ th traffic light. In order to reflect the randomness of the passed green light time in different traffic lights at the start, the initial periodic time  $t_0$  is set to different values for different traffic lights in two driving scenarios.

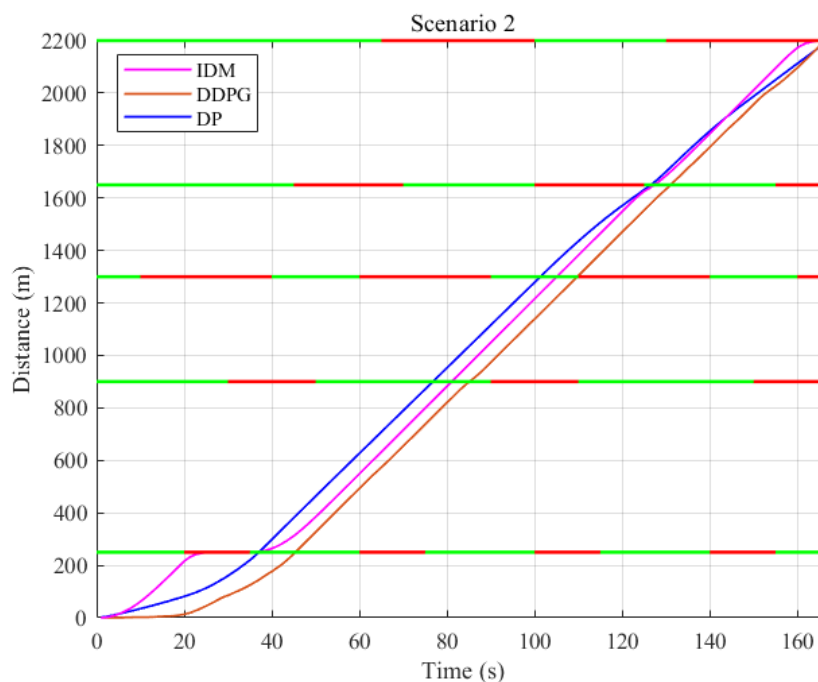
**Table 2.** Parameter settings of road scenario.

|                                 | Scenario 1                             | Scenario 2                 |
|---------------------------------|--|----------------------------|
| Route length $L$ (m)            | 3000                                   | 2200                       |
| Traffic light number $n$ (-)    | 7                                      | 5                          |
| Position $S$ (m)                | 300, 700, 1000, 1700, 2100, 2500, 3000 | 250, 900, 1300, 1650, 2200 |
| Green light duration $T_G$ (s)  | 35, 20, 30, 40, 25, 30, 20             | 25, 40, 20, 30, 30         |
| Red light duration $T_R$ (s)    | 25, 30, 25, 30, 20, 30, 40             | 15, 20, 30, 25, 35         |
| Initial periodic time $t_0$ (s) | 40, 10, 20, 50, 25, 15, 50             | 20, 30, 40, 10, 0          |

Under the principle of complying with traffic rules, in order to improve the traveling efficiency and fuel economy of the HFCHV through traffic light intersection scenarios, this section carries out simulation tests with three different algorithms in two driving test scenarios, and the trajectory of the HFCHV is shown in Figures 4 and 5.



**Figure 4.** Position trajectories of HFCHV in scenario 1. (The red line means Red light during the scenario 1. The green line means Green light during the scenario 1).



**Figure 5.** Position trajectories of HFCHV in scenario 2. (The red line means Red light in the scenario 2. The green line means Green light in the scenario 2).

As shown in Figure 4, in driving scenario 1, the total traveling time of the HFCHV using DP, DDPG and IDM algorithms are 235 s, 240 s and 292 s, respectively. The number of the idling/stop situations of the DP, DDPG and IDM algorithms are 1, 0 and 4, respectively. Compared with the IDM algorithm, the traveling efficiency of the HFCHV using the DDPG algorithm is improved by 17.8% and there is no idling/stop situation. Therefore, in terms of traveling efficiency and driving smoothness, the speed planning of the HFCHV using the DDPG algorithm is significantly better than that applying the IDM algorithm. Meanwhile, the difference in traveling time between the HFCHV using the DDPG algorithm and the HFCHV using the DP algorithm is only 5 s, which is closer to the overall optimized result.

As shown in Figure 5, the total time taken by the HFCHV using the DDPG, DP and IDM algorithms in driving scenario 2 is 165 s, 164 s and 165 s, respectively, and the traveling time is essentially the same. However, the HFCHV using the IDM algorithm has idle stops when passing through the first signal light, while the HFCHV using the DDPG and DP algorithms can avoid this situation, which will favor the smooth driving style adopted by the HFCHV. Therefore, in the upper-level speed planning, the HFCHV using the DDPG algorithm can present superior optimization results in different driving scenarios, and the overall performance is better than that using the IDM algorithm.

## 6.2. Optimization of Hierarchical Eco-Driving Algorithms

Based on the analysis above, the three speed planning algorithms (IDM, DDPG and DP) for the HFCHV passing through traffic lights are established, respectively. In this section, with the aim of minimizing the hydrogen consumption of the hydrogen fuel cell and combining with the optimization results of the upper-level speed planning algorithm, three hierarchical eco-driving optimization algorithms (IDM-DP, DDPG-DP and DP-DP) are established for the HFCHV, and the simulation analysis is carried out in driving scenario 1 and driving scenario 2.

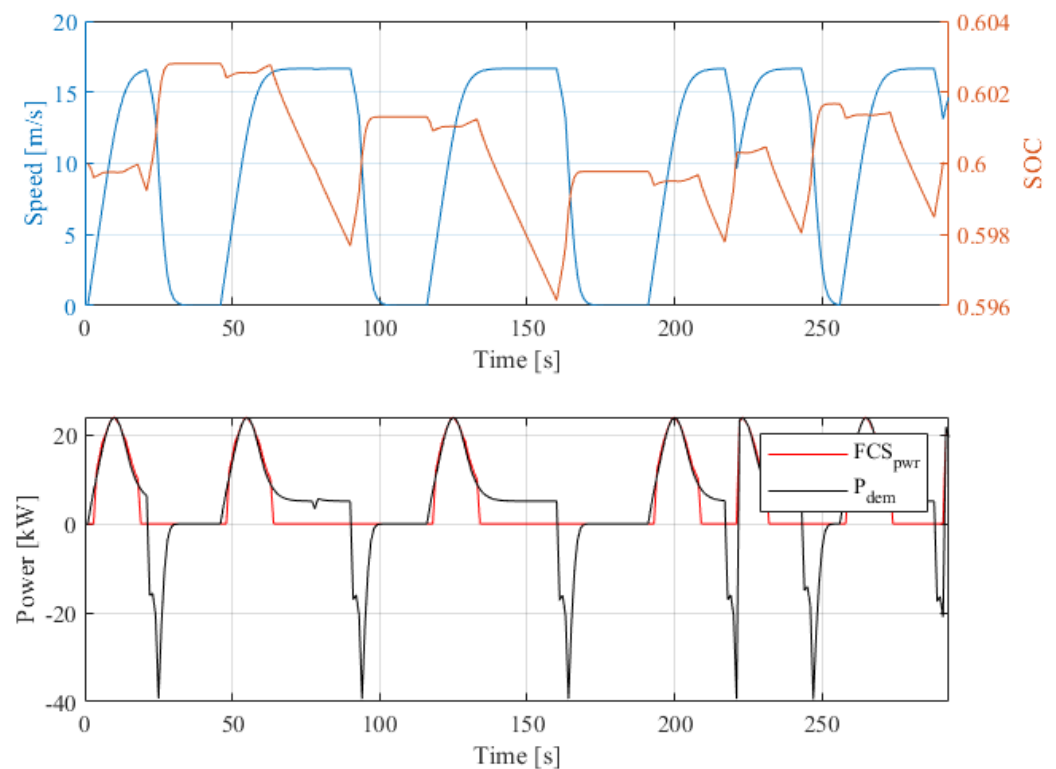
As can be seen from Figures 6–8, during the battery SOC-maintenance simulation process, the battery final SOC based on the three hierarchical algorithms can all be maintained at around 0.6. The driving behaviors of large acceleration and deceleration as well as stopping for the HFCHV using the IDM algorithm during the upper-level speed planning process. Therefore, as shown in Table 3, the HFCHV using the IDM-DP algorithm in driving

scenario 1 has the highest hydrogen consumption of 22.0786 g. Combined with the results in Section 6.1, it can be seen that the traveling efficiency and fuel economy of the HFCHV using the DDPG-DP algorithm is better than that using the IDM-DP algorithm. Meanwhile, the traveling efficiency of the HFCHV using the DDPG-DP algorithm is closer to that using the DP-DP algorithm.

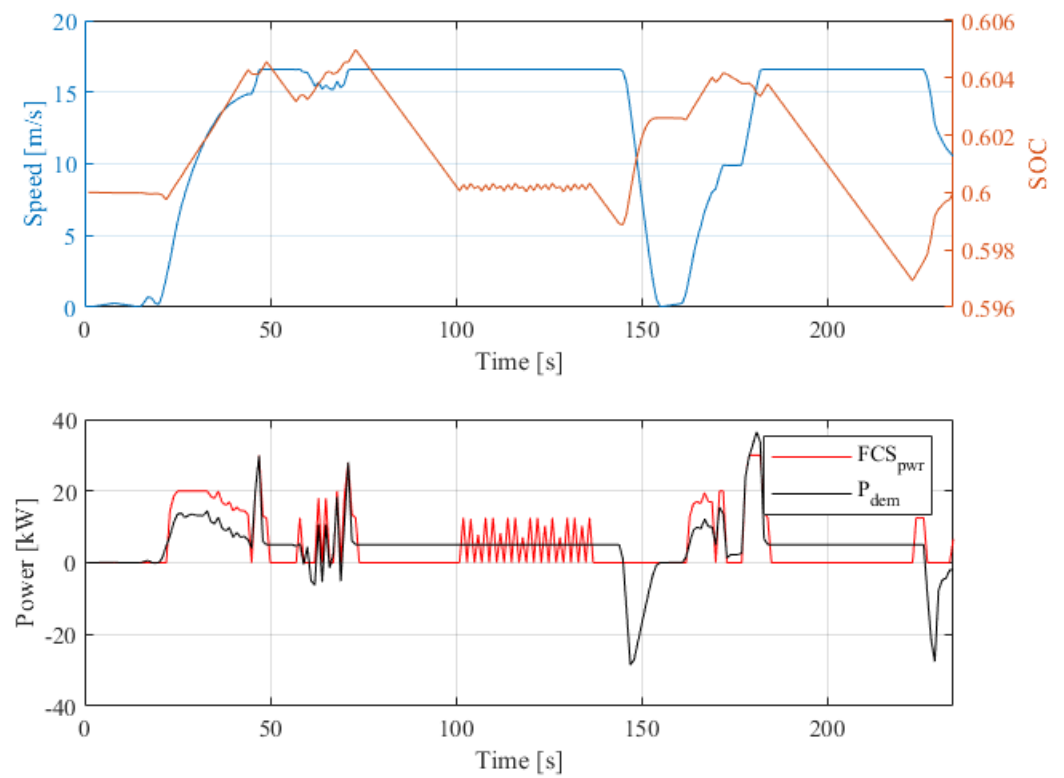
As shown in Figures 9–11, the final SOC of the HFCHV battery using the three algorithms can still be maintained at around 0.6 in scenario 2. Additionally, as can be seen in Figure 9, the overall vehicle power of the HFCHV using the IDM algorithm is significantly increased due to the speed drastic change, which results in a significant increase in the power output of the hydrogen fuel cell. As shown in Table 4, the hydrogen consumption of the HFCHV using the DDPG-DP and IDM-DP algorithms is 13.7714 g and 13.6093 g, respectively, although the hydrogen consumption of the HFCHV using the DDPG-DP algorithm is slightly higher than that using the IDM-DP algorithm. In the upper-level speed planning process, the HFCHV using the DDPG algorithm is able to pass through multiple traffic lights smoothly without stopping, whereas the HFCHV using the IDM algorithm needs to stop and wait, which undoubtedly increases the burden on the road traffic access. As a result, compared with the HFCHV using the IDM-DP algorithm, the HFCHV using the DDPG-DP algorithm is more advantageous in terms of comprehensive performance with almost the same fuel consumption.

**Table 3.** Comparison of hydrogen fuel economy in scenario 1.

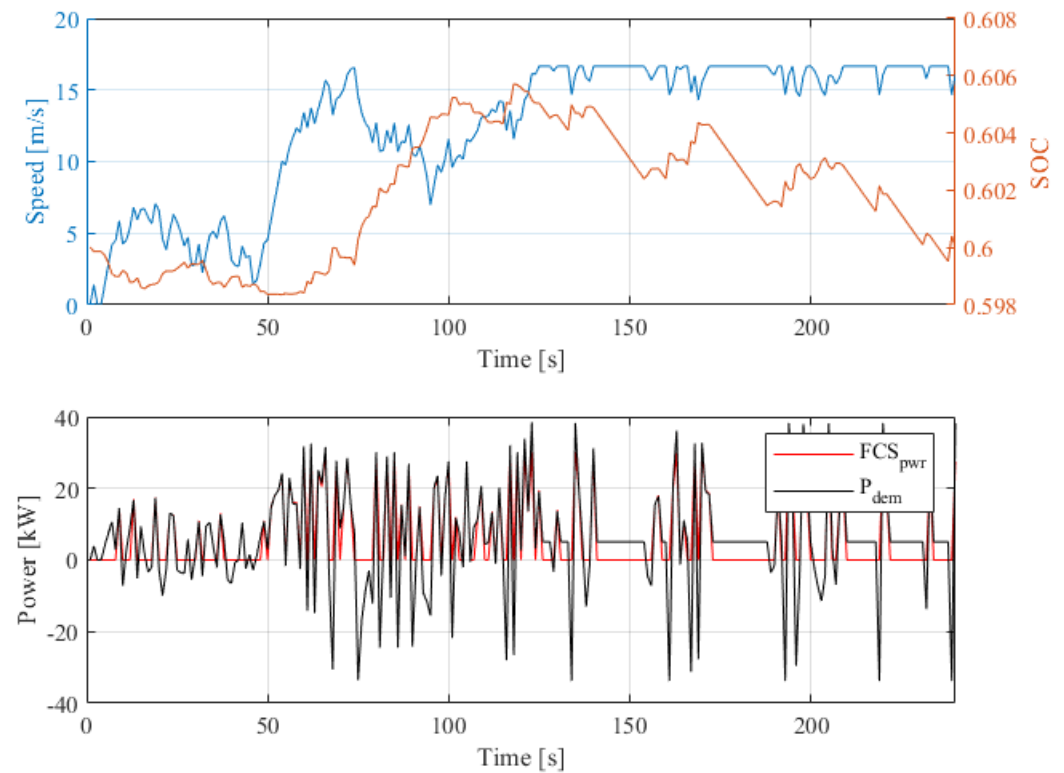
| Hierarchical Algorithms | Hydrogen Consumption (g) | Final SOC |
|-------------------------|--------------------------|-----------|
| DDPG-DP                 | 21.4832                  | 0.601     |
| IDM-DP                  | 22.0786                  | 0.6       |
| DP-DP                   | 16.8230                  | 0.598     |



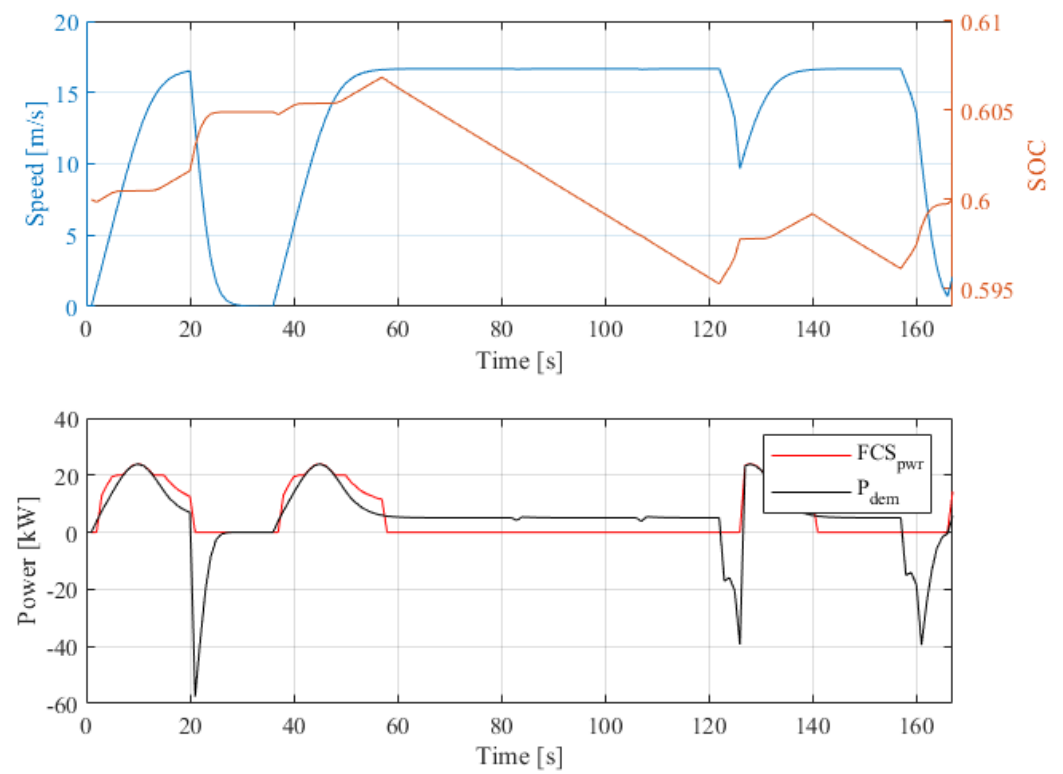
**Figure 6.** Performance of HFCHV using IDM-DP algorithm in scenario 1. (The blue line represents speed of vehicle, the orange line represents SOC of battery).



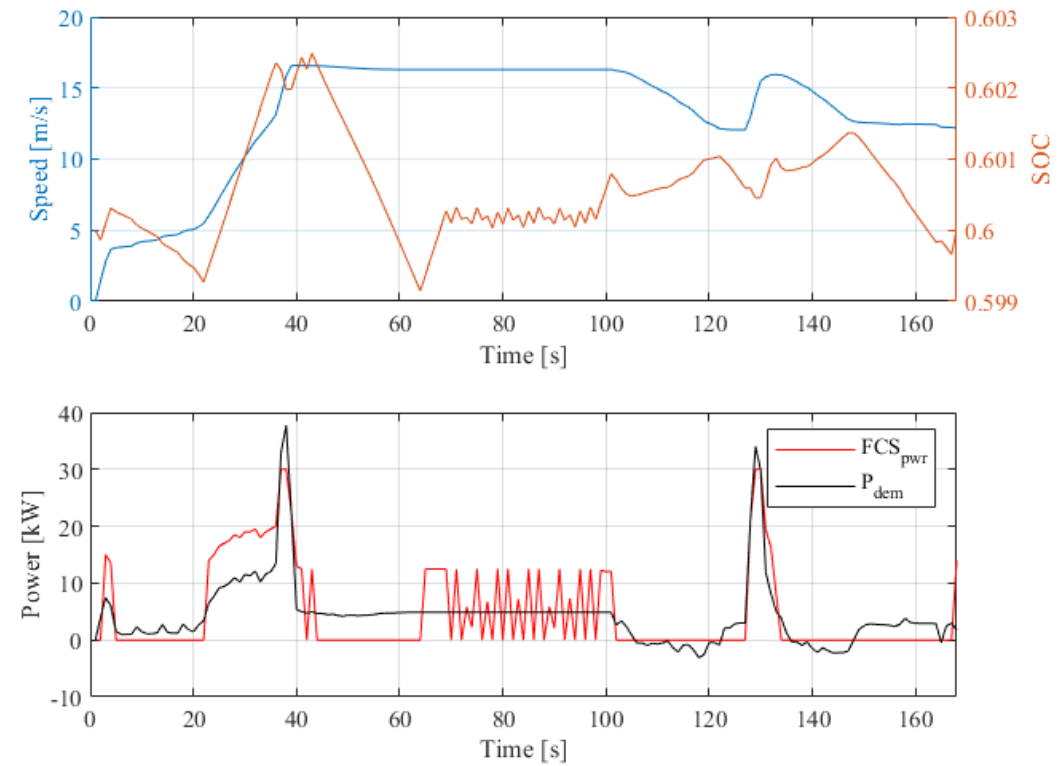
**Figure 7.** Performance of HFCHV using DP-DP algorithm in scenario 1. (The blue line represents speed of vehicle, the orange line represents SOC of battery).



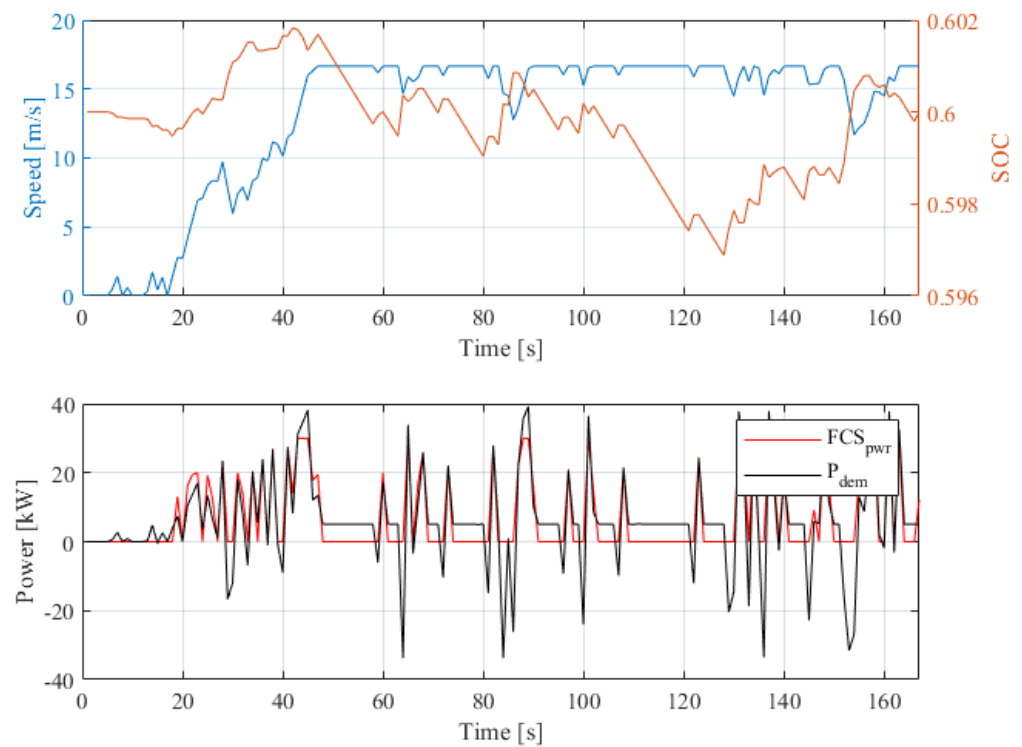
**Figure 8.** Performance of HFCHV using DDPG-DP algorithm in scenario 1. (The blue line represents speed of vehicle, the orange line represents SOC of battery).



**Figure 9.** Performance of HFCHV using IDM-DP algorithm in scenario 2. (The blue line represents speed of vehicle, the orange line represents SOC of battery).



**Figure 10.** Performance of HFCHV using DP-DP algorithm in scenario 2. (The blue line represents speed of vehicle, the orange line represents SOC of battery).



**Figure 11.** Performance of HFCHV using DDPG-DP algorithm in scenario 2. (The blue line represents speed of vehicle, the orange line represents SOC of battery).

**Table 4.** Comparison of fuel consumption in scenario 2.

| Hierarchical Algorithms | Hydrogen Consumption (s) | Final SOC |
|-------------------------|--------------------------|-----------|
| DDPG-DP                 | 13.7714                  | 0.6       |
| IDM-DP                  | 13.6093                  | 0.6       |
| DP-DP                   | 11.3240                  | 0.6       |

## 7. Conclusions

The paper proposes a hierarchical multi-objective optimization algorithm for speed planning and energy management of the HFCHV using the DDPG-DP algorithm. In terms of the upper-level speed planning for the HFCHV, the algorithms of IDM, DDPG and DP are proposed, respectively, and a multi-objective cost function is designed in the DDPG algorithm. For lower-level energy management for the HFCHV, the DP algorithm for energy management is proposed to minimize the overall hydrogen consumption based on the upper-level speed planning algorithm. Meanwhile, two traffic driving scenarios are established in the meantime for verification and the results are shown as follows:

(1) In terms of traveling efficiency, the DDPG algorithm in the HFCHV's upper-level speed planning can significantly reduce idling/stopping, avoid parking and sudden acceleration and deceleration during the driving process in the two driving scenarios. Because this can improve the vehicle traveling efficiency, it is beneficial for reducing the road traffic jam and extending the lifetime of the vehicle.

(2) In terms of fuel economy, in driving scenario 1, the HFCHV using the DDPG-DP algorithm improves by 2.7% compared to that using the IDM-DP algorithm. Although the fuel economy of the HFCHV using the DDPG-DP algorithm is essentially the same as that of the IDM-DP algorithm in driving scenario 2, the HFCHV using the DDPG-DP algorithm reduces the number of idling/stop situation, which is more meaningful for improving road traffic smoothness.

Based on the analysis of the previous result, compared to the IDM-DP algorithm, the proposed DDPG-DP algorithm has a better comprehensive performance in the field of

decreasing fuel consumption and improving the traveling efficiency of the HFCHV. In future work, the energy-saving potential of the HFCHV will be further enhanced by combining speed planning with energy management. Meanwhile, the computation efficiency of the algorithm will be further improved to guarantee the online application of the HFCHV.

**Author Contributions:** Conceptualization, H.D. and L.Z.; methodology, H.D.; software, H.D.; validation, H.D. and L.Z.; formal analysis, H.D. and H.Z.; investigation, H.D. and H.Z.; resources, H.Z. and H.L.; data curation, H.Z. and H.L.; writing—original draft preparation, H.D. and L.Z.; writing—review and editing, H.D. and L.Z.; visualization, H.D. and H.Z.; supervision, H.D. and H.Z.; project administration, H.D. and L.Z.; funding acquisition, H.D., L.Z., H.Z. and H.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Taiyuan University of Science and Technology Scientific Research Initial Funding [20232005].

**Data Availability Statement:** The data are available from the corresponding author on reasonable request.

**Acknowledgments:** This project is supported by Taiyuan University of Science and Technology Scientific Research Initial Funding (20232005). We would like to thank the sponsor.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Zhang, F.; Hu, X.; Langari, R.; Cao, D. Energy management strategies of connected HEVs and PHEVs: Recent progress and outlook. *Prog. Energy Combust. Sci.* **2019**, *73*, 235–256. [[CrossRef](#)]
2. Khan, M.U.; Hosseinzadeh, M.; Mosavi, A. An Intersection-Based Routing Scheme Using Q-Learning in Vehicular Ad Hoc Networks for Traffic Management in the Intelligent Transportation System. *Mathematics* **2022**, *10*, 3731. [[CrossRef](#)]
3. Lv, Z.; Lou, R.; Singh, A.K. AI Empowered Communication Systems for Intelligent Transportation Systems. *IEEE Trans. Intell. Transp. Syst.* **2021**, *22*, 4579–4587. [[CrossRef](#)]
4. Ma, G.Q.; Ghasemi, M.; Song, X.Y. Integrated Powertrain Energy Management and Vehicle Coordination for Multiple Connected Hybrid Electric Vehicles. *IEEE Trans. Veh. Technol.* **2018**, *67*, 2893–2899. [[CrossRef](#)]
5. Zegong, N.; Hongwen, H.; Yong, W.; Ruchen, H. Energy Management Optimization for Connected Hybrid Electric Vehicle with Offline Reinforcement Learning. In Proceedings of the 2022 IEEE 12th International Conference on Electronics Information and Emergency Communication (ICEIEC), Beijing, China, 15–17 July 2022; pp. 103–106.
6. Shen, D.; Lim, C.-C.; Shi, P. Fuzzy Model Based Control for Energy Management and Optimization in Fuel Cell Vehicles. *IEEE Trans. Veh. Technol.* **2020**, *69*, 14674–14688. [[CrossRef](#)]
7. Zou, W.; Li, J.; Yang, Q.; Duan, Z. An Improved Max-Min Game Theory Control of Fuel Cell and Battery Hybrid Energy System against System Uncertainty. *IEEE J. Emerg. Sel. Top. Power Electron.* **2022**, *11*, 78–87. [[CrossRef](#)]
8. Huang, Y.; Hu, H.; Tan, J.; Lu, C.; Xuan, D. Deep reinforcement learning based energy management strategy for range extend fuel cell hybrid electric vehicle. *Energy Convers. Manag.* **2023**, *277*, 116678. [[CrossRef](#)]
9. Huo, W.; Chen, D.; Tian, S.; Li, J.; Zhao, T.; Liu, B. Lifespan-consciousness and minimum-consumption coupled energy management strategy for fuel cell hybrid vehicles via deep reinforcement learning. *Int. J. Hydrogen Energy* **2022**, *47*, 24026–24041. [[CrossRef](#)]
10. Peng, J.K.; He, H.W.; Xiong, R. Rule based energy management strategy for a series-parallel plug-in hybrid electric bus optimized by dynamic programming. *Appl. Energy* **2017**, *185*, 1633–1643. [[CrossRef](#)]
11. Castaigns, A.; Lhomme, W.; Trigui, R.; Bouscayrol, A. Comparison of energy management strategies of a battery/supercapacitors system for electric vehicle under real-time constraints. *Appl. Energy* **2016**, *163*, 190–200. [[CrossRef](#)]
12. Xie, S.B.; Hu, X.S.; Xin, Z.K.; Brighton, J. Pontryagin's Minimum Principle based model predictive control of energy management for a plug-in hybrid electric bus. *Appl. Energy* **2019**, *236*, 893–905. [[CrossRef](#)]
13. Sun, C.; Sun, F.C.; He, H.W. Investigating adaptive-ECMS with velocity forecast ability for hybrid electric vehicles. *Appl. Energy* **2017**, *185*, 1644–1653. [[CrossRef](#)]
14. Du, G.; Zou, Y.; Zhang, X.; Guo, L.; Guo, N. Energy management for a hybrid electric vehicle based on prioritized deep reinforcement learning framework. *Energy* **2022**, *241*, 122523. [[CrossRef](#)]
15. Wu, Y.; Tan, H.; Peng, J.; Zhang, H.; He, H. Deep reinforcement learning of energy management with continuous control strategy and traffic information for a series-parallel plug-in hybrid electric bus. *Appl. Energy* **2019**, *247*, 454–466. [[CrossRef](#)]
16. Chen, Z.; Mi, C.C.; Xu, J.; Gong, X.; You, C. Energy Management for a Power-Split Plug-in Hybrid Electric Vehicle Based on Dynamic Programming and Neural Networks. *IEEE Trans. Veh. Technol.* **2014**, *63*, 1567–1580. [[CrossRef](#)]
17. Jia, C.; Zhou, J.; He, H.; Li, J.; Wei, Z.; Li, K.; Shi, M. A novel energy management strategy for hybrid electric bus with fuel cell health and battery thermal- and health-constrained awareness. *Energy* **2023**, *271*, 127105. [[CrossRef](#)]

18. Tang, X.L.; Zhou, H.T.; Wang, F.; Wang, W.D.; Lin, X.K. Longevity-conscious energy management strategy of fuel cell hybrid electric Vehicle Based on deep reinforcement learning. *Energy* **2022**, *238*, 121593. [[CrossRef](#)]
19. Zhou, J.; Feng, C.; Su, Q.; Jiang, S.; Fan, Z.; Ruan, J.; Sun, S.; Hu, L. The Multi-Objective Optimization of Powertrain Design and Energy Management Strategy for Fuel Cell–Battery Electric Vehicle. *Sustainability* **2022**, *14*, 6320. [[CrossRef](#)]
20. Zhang, Q.; Ju, F.; Zhang, S.; Deng, W.; Wu, J.; Gao, C. Power Management for Hybrid Energy Storage System of Electric Vehicles Considering Inaccurate Terrain Information. *IEEE Trans. Autom. Sci. Eng.* **2017**, *14*, 608–618. [[CrossRef](#)]
21. Yang, Y.; Su, L.; Qin, D.; Gong, H.; Zeng, J. Energy management strategy for hybrid electric vehicle based on system efficiency and battery life optimization. *Wuhan Univ. J. Nat. Sci. China* **2014**, *19*, 269–276. [[CrossRef](#)]
22. Jia, C.; Li, K.; He, H.; Zhou, J.; Li, J.; Wei, Z. Health-aware energy management strategy for fuel cell hybrid bus considering air-conditioning control based on TD3 algorithm. *Energy* **2023**, *283*, 128462. [[CrossRef](#)]
23. Anselma, P.G.; Kollmeyer, P.; Lempert, J.; Zhao, Z.; Belingardi, G.; Emadi, A. Battery state-of-health sensitive energy management of hybrid electric vehicles: Lifetime prediction and ageing experimental validation. *Appl. Energy* **2021**, *285*, 116440. [[CrossRef](#)]
24. Feng, Y.B.; Dong, Z.M. Optimal energy management strategy of fuel-cell battery hybrid electric mining truck to achieve minimum lifecycle operation costs. *Int. J. Energy Res.* **2020**, *44*, 10797–10808. [[CrossRef](#)]
25. Asadi, B.; Vahidi, A. Predictive Cruise Control: Utilizing Upcoming Traffic Signal Information for Improving Fuel Economy and Reducing Trip Time. *IEEE Trans. Control Syst. Technol.* **2011**, *19*, 707–714. [[CrossRef](#)]
26. Wang, Z.; Wu, G.; Barth, M.J. Cooperative Eco-Driving at Signalized Intersections in a Partially Connected and Automated Vehicle Environment. *IEEE Trans. Intell. Transp. Syst.* **2020**, *21*, 2029–2038. [[CrossRef](#)]
27. Zhang, R.; Yao, E. Eco-driving at signalised intersections for electric vehicles. *IET Intell. Transp. Syst.* **2015**, *9*, 488–497. [[CrossRef](#)]
28. Hu, J.; Shao, Y.; Sun, Z.; Wang, M.; Bared, J.; Huang, P. Integrated optimal eco-driving on rolling terrain for hybrid electric vehicle with vehicle-infrastructure communication. *Transp. Res. Part C Emerg. Technol.* **2016**, *68*, 228–244. [[CrossRef](#)]
29. Kim, Y.; Figueroa-Santos, M.; Prakash, N.; Baek, S.; Siegel, J.B.; Rizzo, D.M. Co-optimization of speed trajectory and power management for a fuel-cell/battery electric vehicle. *Appl. Energy* **2020**, *260*, 114254. [[CrossRef](#)]
30. Liu, Y.G.; Huang, Z.Z.; Li, J.; Ye, M.; Zhang, Y.J.; Chen, Z. Cooperative optimization of velocity planning and energy management for connected plug-in hybrid electric vehicles. *Appl. Math. Model.* **2021**, *95*, 715–733. [[CrossRef](#)]
31. Bai, Z.; Hao, P.; Shangguan, W.; Cai, B.; Barth, M.J. Hybrid Reinforcement Learning-Based Eco-Driving Strategy for Connected and Automated Vehicles at Signalized Intersections. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 15850–15863. [[CrossRef](#)]
32. Wang, Q.Z.; Gong, Y.B.; Yang, X.F. Connected automated vehicle trajectory optimization along signalized arterial: A decentralized approach under mixed traffic environment. *Transp. Res. Part C Emerg. Technol.* **2022**, *145*, 103918. [[CrossRef](#)]
33. Liu, B.; Sun, C.; Wang, B.; Liang, W.; Ren, Q.; Li, J.; Sun, F. Bi-level convex optimization of eco-driving for connected Fuel Cell Hybrid Electric Vehicles through signalized intersections. *Energy* **2022**, *252*, 123956. [[CrossRef](#)]
34. Dong, H.; Zhuang, W.; Chen, B.; Lu, Y.; Liu, S.; Xu, L.; Pi, D.; Yin, G. Predictive energy-efficient driving strategy design of connected electric vehicle among multiple signalized intersections. *Transp. Res. Part C Emerg. Technol.* **2022**, *137*, 103595. [[CrossRef](#)]

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