IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 4, 2019, accepted December 29, 2019, date of publication January 6, 2020, date of current version January 16, 2020.

*Digital Object Identifier 10.1109/ACCESS.2020.2964116*

# Semi-Active Suspension Control Based on Deep Reinforcement Learning

**LIU MING[1,2], LI YIBIN[1], RONG XUEWEN[1], ZHANG SHUAISHUAI[2], AND YIN YANFANG[2]**
[1]College of Control Science and Engineering, Shandong University, Jinan 250061, China
[2]Department of Electrical Information, Shandong University of Science and Technology, Jinan 250031, China

Corresponding author: Li Yibin (liyb@sdu.edu.cn)

This work was supported in part by the General Program of National Natural Science Foundation of China under Grant U1613223 and Grant 61703243, and in part by the Major Program of Shandong Province Natural Science Foundation under Grant ZR2018ZC0437.

**ABSTRACT** The performance of vehicle body vibration and ride comfort of active or semi-active suspension with proper control is better than that with passive suspension. The key to achieve good control effect is that the suspension control system should have strong real-time learning ability according to changes in the road surface and suspension parameters. In the control strategies adopted by previous researchers, the classical neural network controller has some learning ability, but it is mainly based on offline learning with a large number of samples. In this paper, the deep reinforcement learning strategy is used to solve the above problems .Aiming at the continuity of state space and execution action in vehicle active suspension system, the control of the semi-active suspension is realized by using improved DDPG (Deep Deterministic Policy Gradient) algorithm. To overcome the shortcoming of low efficiency of this algorithm in the initial stage of learning, the DDPG algorithm is improved and using empirical samples in the learning method is proposed. Based on Mujoco, the physical model of semi-active suspension is established, and its dynamic characteristics are analyzed under the condition of various road level and vehicle speed. The simulation results show that compared with the passive suspension, the semi-active suspension based on improved DDPG algorithm with learning method using experienced samples can better adapt to various road level, more effectively reduce the vertical acceleration of the vehicle body and the dynamic deflection of the suspension, and further improve the ride comfort.

**INDEX TERMS** Semi-active suspension, deep reinforcement learning, DDPG, experienced samples.

## I. INTRODUCTION

As one of the important part of vehicles, suspension system has a huge impact on vehicle handling stability, driving safety and riding comfort. According to the working principle, suspension system can be summarized as passive suspension, semi-active suspension and active suspension. Since the parameters of passive suspension are fixed and unchanged, when the driving environment or vehicle parameters change, the ride comfort, stability and comfort of the vehicle can't be guaranteed. Therefore, semi-active suspension and active suspension with adjustable dynamic parameters, which can overcome the above shortcomings of passive suspension, have been the research hotspot in recent decades.

In practical vehicle suspension systems, the parameter uncertainties are usually unavoidable. To solve the problem, researchers have studied different modern control strategies. In addition to the traditional PID control strategy, they also adopt control strategies such as SDC (Skyhook Damper Control) [1], LQG (Linear Quadratic Gaussian) control [2], [3], fuzzy control [4], [5], SMC (Sliding Mode Control) [6], $H\infty$ control [7], [8], adaptive control [9], [10], predictive control [11], [12], neural network control [13], [14], nonlinear intelligent control [15], and compound control [16], to reduce the vibration amplitude of vehicle body and improve the vehicle ride smoothness, stability, and ride comfort.

In the previous studies of active suspension control, PID control is often use as it has a simple principle and strong adaptability. However, it is not very effective in suspension systems with uncertain parameters. The skyhook control in [1] is a classic semi-active suspension control method.

The associate editor coordinating the review of this manuscript and approving it for publication was Sudhakar Babu Thanikanti.

It has the characteristics of simple structure, reliable performance and easy engineering implementation. However, the time delay of the system has an obvious impact on the suspension performance, which will lead to the instability of the suspension system and wheel jump. Using half vehicle suspension model, Lan and Yu [2] designed the LQG controller of active suspension by using optimal control theory and its control effect was verified by simulation. Yu *et al.* [3] presented a method to determine the optimal control weighting coefficient and optimal control force in the design of LQG controller. But the control parameters are calculated on the premise of ignoring the uncertain factors in the system modeling. Once the system parameters change to a certain extent, it will make the system become unstable. In [4], the vehicle acceleration is used as the input of fuzzy controller to design the control rules and effectiveness is verified. The topic of [5] is to solve the problem of blind design of fuzzy control rules, but it's actually PID control. In fact, Fuzzy control rule development needs a lot of control experience and knowledge, and the control effectiveness is closely related to the rules defined. Sliding mode control strategy in [6] shows good control effect and robustness to the control of active suspension system. However, chattering is the fatal disadvantage of sliding mode control, which sometimes leads to instability or even unavailability of the control system. Du and Zhang [7] designed a constrained delay-dependent $H\infty$ state feedback controller and verified the effectiveness in spite of the existence of a time delay in control input. Sun *et al.* [8] used the generalized Kalman-Yakubovich-Popov (KYP) lemma to reduce the norm of disturbance to the control output in a specific frequency band, and a state feedback controller was designed under the LMI optimization framework. However, good robustness comes at the expense of other performance of the system. The essence of the adaptive control strategy adopted in [9], [10] is to ensure the performance of the system by detecting the parameter changes of the suspension system to determine the control parameters. However, the adaptive control strategy can only be adjusted within a certain range, once it exceeds the range, the control effect will become worse. Milanese *et al.* [11] presented a fast model predictive control (FMPC) implementation method based on nonlinear function approximation technology to solve the problem of fast calculation of predictive control rate, and its effectiveness is verified by simulation. Wang *et al.* [12] designed a robust model predictive controller (RMPC) for active suspension system by defining the performance evaluation function of RMPC. The accuracy and effectiveness of the controller were verified by prototype vehicle simulation and road test. The predictive control methods used in these two papers transmit the road information to the suspension device in advance. The key is to obtain the information with certain accuracy, undisturbed and reflecting the real situation of the road. But it is almost impossible to meet the above conditions. The neural network control method designed in [13], [14] makes full use of the fact that neural networks are suitable for nonlinear system, control and achieve good results. However, as a supervised learning method, it requires the system provide many samples with labels. In the process of generating the control strategy, only the current state is considered, but not the future state, which severely limits the utility of the method. Qin *et al.* [15] presented a new adaptive nonlinear intelligent control strategy based on road classification, and the effectiveness of the method was verified by simulation with different road levels. However, due to the use of data-driven road classification method, when the control strategy changes, it is necessary to reconstruct a new classifier. Sharma *et al.* [16] used disturbance rejection and continuous state damper controllers to reduce lateral vibration of passenger rail vehicle and verified the effectiveness by simulation. However, only specific disturbances are considered in system design and this approach has no universal applicability.

Considering the characteristics of suspension system parameters variation and uncertainty, the control system should have continuous learning functions. That means the system can continuously improve the control force of the active suspension system by its own learning mechanism according to the actual road and suspension parameters.

DRL (Deep Reinforcement Learning) is one of the new research hotspots in the field of machine learning. It combines the perception ability of DL (Deep Learning) with the decision-making ability of RL (Reinforcement Learning). It can be used to realize the direct control from the original input to the output through the end-to-end way, which has a strong versatility. DRL has been widely used in the field of electronic game since it was put forward. At the same time, it has been successfully applied in parameter optimization, robot control, machine vision, natural language processing and other fields [17].

At present, the main DRL methods include DRL based on value function, DRL based on policy gradient, and deep DRL based on search and supervision. DRL algorithms mainly include DQN (Deep Q-Network) algorithm, DDPG algorithm, A3C (Asynchronous Advantage Actor-Critic) algorithm and DRL method based on MCST (Monte Carlo Tree Search). Alpha Go, the most famous robot designed by Deepmind, uses DRL algorithm based on MCST [18]. The DQN algorithm proposed by this company can learn to play games by directly observing Atari 2600's game pictures and scoring information, and it is universal for almost all games.

The learning mechanism of DRL is learning while obtaining samples. This method updates its model after obtaining the sample and uses the current model to guide the next action, so it can generalize a complex situation that it has never experienced before. This is exactly what is needed for active suspension control in the face of different road conditions and other uncertainties.

To solve the problem of the ride comfort and stability of the active suspension control system during the learning process, the passive suspension is still retained in the design of the suspension system. In order to make full use of the advantages of passive and active suspension force control and its practical application, a force control strategy for vehicle

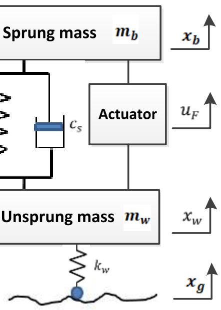



**FIGURE 1.** **Quarter vehicle model with semi-active suspension.**

semi-active suspension system based on deep reinforcement learning is proposed. Considering that the force control of the suspension system of the intelligent integrated electric wheel is a continuous action, the DDPG algorithm is adopted in semi-active suspension control system. This method can automatically adjust the parameters according to the external changes of the unknown model, to achieve the optimal control of vehicle ride comfort and stability. In order to improve the learning speed of the model, the training approach of the DDPG algorithm is improved by using some good control experience samples as the reference.

## II. QUARTER VEHICLE MODEL WITH SEMI-ACTIVE SUSPENSION SYSTEM

The dynamic model of vehicle active or semi-active suspension system includes the quarter vehicle model, half vehicle model and the whole vehicle model. The control algorithm based on the multi-degree-of-freedom model of the vehicle is complex and requires many sensors for collecting a large amount of data, so it is difficult to apply in practice. Therefore, considering the vertical motion of semi-active suspension body and non-spring mass, a 2-DOF quarter vehicle dynamic model is established in this paper, which is commonly used in prior work as shown in figure 1[19].

In Figure 1, $m_b$ is the mass of the vehicle body, $k_s$ is the spring stiffness, $c_s$ is the damping coefficient, $m_w$ is the mass of the wheel axle, $k_w$ is the equivalent stiffness of the tire, $u_F$ is the active control force generated by the linear motor actuator, $x_g$ is the road surface excitation, $x_b$ and $x_w$ are the absolute displacement of the body and tire respectively, $x_b - x_w$ is the suspension deflection, $\ddot{x_b}$ is the body acceleration.

According to the dynamic model shown in Fig. 1, using Newton's second law, the differential equation of longitudinal motion of the body and the wheel can be obtained as:

$$\begin{cases} m_w\ddot{x_w} = k_s\left(x_b - x_w\right) + c_s\left(\dot{x_b} - \dot{x_w}\right) - u_F \\ \qquad\qquad + k_w\left(x_w - x_g\right) \\ m_b\ddot{x_b} = -k_s\left(x_b - x_w\right) - c_s\left(\dot{x_b} - \dot{x_w}\right) + u_F \end{cases} \tag{1}$$

Define the state variables as:

$$\begin{aligned} x_1 &= x_w, \quad x_2 = x_b, \; x_3 = \dot{x_w}, \; x_4 = \dot{x_b} \\ u_1 &= x_g, \quad u_2 = u_F \\ y_1 &= \ddot{x_b}, \quad y_2 = x_b - x_w \end{aligned}$$

Formula (1) can be rewritten as the following state-space form:

$$\begin{cases} \dot{X} = AX + BU \\ Y = CX + DU \end{cases} \tag{2}$$

where,

$$\begin{aligned} X &= \begin{bmatrix} x_w & x_b & \dot{x_w} & \dot{x_b} \end{bmatrix}^T \\ U &= \begin{bmatrix} x_g & u_F \end{bmatrix}^T \\ Y &= \begin{bmatrix} \ddot{x_b} & x_b - x_w \end{bmatrix}^T \\ A &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\dfrac{k_s + k_w}{m_w} & \dfrac{k_s}{m_w} & -\dfrac{c_s}{m_w} & \dfrac{c_s}{m_w} \\ \dfrac{k_s}{m_b} & -\dfrac{k_s}{m_b} & \dfrac{c_s}{m_b} & -\dfrac{c_s}{m_w} \end{bmatrix} \\ B &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ -\dfrac{k_w}{m_w} & -\dfrac{1}{m_w} \\ 0 & \dfrac{1}{m_b} \end{bmatrix} \\ C &= \begin{bmatrix} k_s & -k_s & c_s & -c_s \\ -1 & 1 & 0 & 0 \end{bmatrix} \\ D &= \begin{bmatrix} 0 & \frac{1}{m_b} \\ 0 & 0 \end{bmatrix} \end{aligned}$$

## III. ROAD ROUGHNESS MODEL

Road roughness model is very important for analyzing the characteristics of suspension system. Generally, road surface roughness is random, and its statistical characteristics can be described by PSD (Power Spectral Density). According to ISO/TC108/SC2N67 and GB7031, the power spectrum of road roughness is fitted by an exponential function as follows:

$$G_q\left(n\right) = G_q\left(n_0\right)\left(\frac{n}{n_0}\right)^{-\omega} \tag{3}$$

where $n_0$ is the reference spatial frequency, generally $n_0 = 0.1m^{-1}$, $G_q\left(n_0\right)$ represents the pavement flatness coefficient ($m^3$), $\omega$ is the frequency index, generally $\omega = 2$[20].

There are many time domain models of road roughness excitation [21]. Among them, the single wheel road excitation model based on the filter white noise method is widely used. The model expression is:

$$\dot{g}\left(t\right) = -2\pi n_{00}v \cdot g\left(t\right) + 2\pi n_0\sqrt{G_q\left(n_0\right)v} \cdot \omega(t) \tag{4}$$

This model is in good agreement with the standard pavement spectrum. The physical meaning of the model is clear, and it can be used as the input excitation for vehicle ride comfort analysis [22].

## IV. SEMI-ACTIVE SUSPENSION CONTROL BASED ON DDPG ALGORITHM

The structure of the semi-active suspension control system based on DDPG is shown in the figure 2.

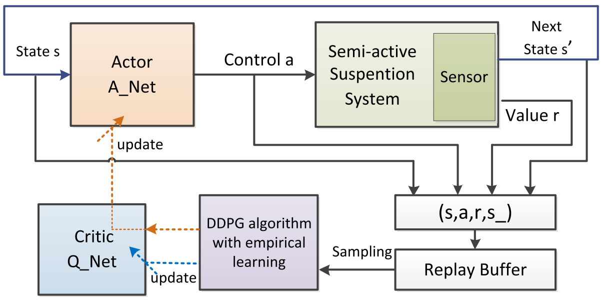



**FIGURE 2. The control structure of semi-active suspension system based on DDPG.**

In this system, there are two tunable networks: actor network and critical network. Actor network obtains the control quantity according to the output state information of the suspension system and acts on it. The system synthesizes the previous and current state information and control information of suspension output, and uses DDPG algorithm to realize the parameter learning and adjustment of actor network and critical network based on the evaluation target.

From formula (1), adjusting the active control force $u_F$ (abbreviated as $u$) of the linear motor actuator will affect the acceleration of the body and the dynamic deflection of the suspension at the current moment, and produce a new state. The new state will also affect the acceleration of the body and the dynamic deflection of the suspension at the next moment.

According to Bellman's formula, the behavior value function of $u_t$ to take action on state $s_t$ under specific control strategy $\pi$ can be expressed as follows:

$$Q^{\pi}\left(s_t, u_t\right) = \mathrm{E}[\mathrm{r}\left(s_t, u_t\right) + \gamma Q^{\pi}\left(s_{t+1}, u_{t+1}\right) \tag{5}$$

where, $u_{t+1}$ is the control quantity adopted by current strategy $\pi$ to the state $s_{t+1}$, $Q^{\pi}\left(s_{t+1}, u_{t+1}\right)$ is the behavior value function of the next moment, and gamma is the discount factor of future returns. The stability evaluation value $\mathrm{r}\left(s_t, u_t\right)$ obtained by feedback from semi-active suspension system after $u_t$ control is performed under the state $s_t$ is defined as:

$$r = -(k_1 y_1^2 + k_2 y_2^2 + k_3 y_3^2) \tag{6}$$

where, $y_1$ is the vehicle body acceleration, $y_2$ is the suspension dynamic deflection, $y_3$ is the vehicle body displacement; $k_1, k_2, k_3$ are the weights of the vehicle body acceleration, the suspension dynamic deflection and the vehicle body displacement in the stability evaluation.

## A. NETWORK MODEL

The $Q$ function calculation in equation (5) is a recursive formula. In the actual situation, it is impossible to recursively calculate the value at each step. Using the A-C architecture [23], the $Q$ value is simulated by the neural network model $Q\_net$ , which is expressed as:

$$Q^{\pi}\left(s_t, u\right) = Q\_net(s_t, u) \tag{7}$$

Semi-active suspension system is a continuous control system. Actions $u$ are executed on state $s_t$ under current policy $\pi$.

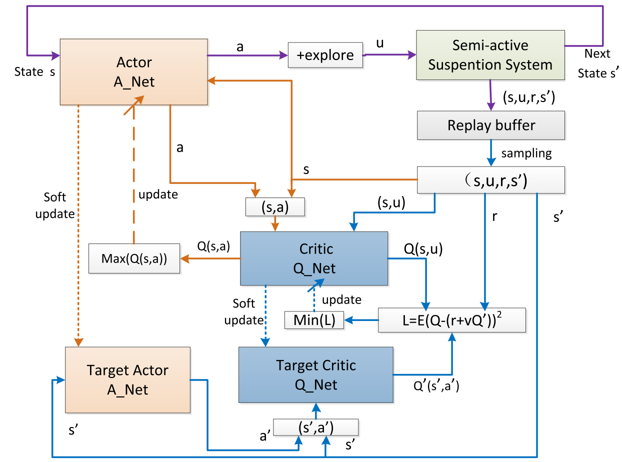



**FIGURE 3. The DDPG algorithm structure.**

It is provided by a policy network model $A\_net$, which is expressed as:

$$u_t = (s_t) \tag{8}$$

The goal of the training evaluation network $Q\_net$ is to minimize the error of the $Q$ value, and the goal of the training strategy network A_*net* is to maximize $Q$ value by output actions. The target functions are as follows:

$$\begin{cases} L\left(\theta^{Q}\right) = \left(Q\_net\left(s_t, u\right) - y_t\right)^2 \\ y_t = r_t + \gamma Q\_net\left(s_{t+1}, A\_net(s_{t+1})\right) \end{cases} \tag{9}$$

$$L\left(\theta^{A}\right) = -Q\_net\left(s_t, A\_net(s_t)\right) \tag{10}$$

## B. DDPG ARCHITECTURE

In the learning process, since the parameters of $Q\_net$ are used to calculate the gradient of $Q\_net$ and $A\_net$ while the gradient is updated, the learning process is more likely to diverge. In order to solve the above problems, the DDPG architecture algorithm is proposed in document [24]. The algorithm creates a replica network of evaluation network $Q\_net$ and policy network $A\_net$ to calculate the gradient respectively. At the same time, the algorithm also creates a finite size replay buffer to store exploration samples and updates the network by uniformly sampling a small batch. To solve the problem of sample independence, the specific architecture is shown in figure 3.

By minimizing the objective functions $L\left(\theta^{Q}\right)$ and $L\left(\theta^{A}\right)$, we can find that the updating network is also used to calculate the objective function of $Q$, so the updating of network learning is difficult to converge. This document put forward the DDPG architecture algorithm, create the replica network of evaluation network $Q\_net$ and policy network $A\_net$ respectively to solve this problem by calculating the objective function. At the same time, a finite size replay buffer is created to store the exploratory samples, and the problem of sample independence is solved by uniformly sampling a

**TABLE 1. DDPG learning algorithm with experience values.**

Procedure: DDPG algorithm with empirical learning

Randomly initialize critic network $\boldsymbol{Q\_net}$ and actor network $\boldsymbol{A\_net}$ with weights $\boldsymbol{\theta}^{A}$ and $\boldsymbol{\theta}^{Q}$

Initialize target network parameters, $\boldsymbol{\theta}^{Q'} \leftarrow \boldsymbol{\theta}^{Q}$, $\boldsymbol{\theta}^{A'} \leftarrow \boldsymbol{\theta}^{A}$

Initialize the buffer

Generate empirical samples $(\mathbf{s_i}, \mathbf{a_i}, \mathbf{r_i} \mathbf{s_{i+1}})$ to fill the buffer

Initialize $\aleph$ for action exploration

for episode=1,..M do

  Receive the initial observation state s1 of env

  for t=1,T do

    $\boldsymbol{u}_t \leftarrow \mathbf{A\_net}(\mathbf{s_t}|\boldsymbol{\theta}^{A})$

      $\mathbf{a_t} \leftarrow \boldsymbol{u}_t + \aleph * \mathbf{random}$

      $\mathbf{r_t}, \mathbf{s_{t+1}} \leftarrow \mathbf{env.step}(\mathbf{a_t})$

    Store samples $(\mathbf{s_t}, \mathbf{a_t}, \mathbf{r_t}, \mathbf{s_{t+1}})$ in buffer pool $\mathbf{R}$

    Sample N samples $(\mathbf{s_i}, \mathbf{a_i}, \mathbf{r_i}, \mathbf{s_{i+1}})$ for R

      $\boldsymbol{y_i = r_i + \gamma Q\_net'(s_{i+1}, A\_net'(s_{i+1}|, \theta^{A'})|\theta^{Q'})}$

      $\boldsymbol{L1 = \frac{1}{N}\sum_i (Q\_net(s_i, a_i|\theta^{Q}) - y_i)^2}$

      Update $\boldsymbol{\theta}^{Q}$ by minimizing $\boldsymbol{L1}$

      $\boldsymbol{L2 = -\frac{1}{N}\sum_i Q\_net(s_t, A\_net(s_t|\theta^{A})|\theta^{Q})}$

      Update $\boldsymbol{\theta}^{Q}$ by minimizing $\boldsymbol{L2}$

    Update the target networks:

$$\boldsymbol{\theta^{A'} \leftarrow \tau|\theta^{A} + (1-\tau)\theta^{A'}}$$

$$\boldsymbol{\theta^{Q'} \leftarrow \tau|\theta^{Q} + (1-\tau)\theta^{Q'}}$$

    end for

  end for

end for

small batch of networks updates. Its architecture is shown in figure 3.

### *C. DDPG ALGORITHM WITH EXPERIENCE LEARNING*

In order to improve the learning and training speed of the model, we consider using some good control experience samples of passive suspension to let the model learn from experience samples. This can provide a good reference for learning and improve the training speed. In this paper, additional forces obtained by increasing or reducing the elastic coefficient of passive suspension are used as control force samples to fill in the initial buffer pool, which are the initial learning experience. Based on this idea, the model learning algorithm is shown in Table 1.

## V. SIMULATION EXPERIMENTS

### *A. BUILDING OF SIMULATION ENVIRONMENT AND PARAMETERS SETTING*

In this paper, MuJoCo is used to build the quarter semi-active suspension reinforcement learning simulation system. Figure 4 is the simulation model of the system designed by MuJoco.

The parameters of suspension model and DDPG model need to be set in the simulation experiment.

The suspension model parameters are shown in Table 2 [25].

The control force is important for the performance of the semi-active suspension. Therefore, we carry out the suspension parameter performance simulation experiment by

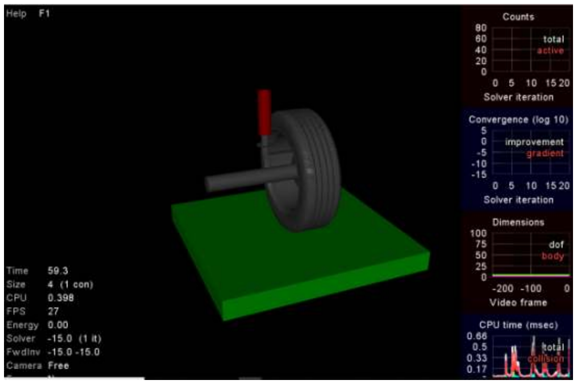


**FIGURE 4. Simulation environment built by Mujoco.**

**TABLE 2. Parameters of vehicle suspension systems.**

| Parameters | Units | Value |
|---|---|---|
| Sprung mass $m_b$ | [kg] | 400 |
| Unsprung mass $m_w$ | [kg] | 40 |
| Suspension stiffness $k_s$ | [N/m] | 21000 |
| Suspension dumping $c_s$ | [N.s/m] | 1500 |
| Tire stiffness $k_w$ | [N/m] | 150000 |

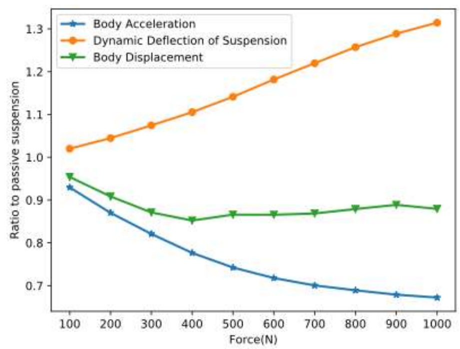



**FIGURE 5. Simulation experiment results by changing the force control range.**

changing the force control range on the C level road with the speed of 60 km/h. Figure 5 demonstrates the simulation results. Considering the comprehensive control performance of semi-active suspension, the force control range is set from – 600N to 600N.

The Critic and the Actor are both neural network with three hidden layers. Dropout is set for each layer, and the rate coefficient are both 0.1.The number of hidden layer nodes of the Critic are 100,100 and 50, and the activation functions are relu, sigmoid and sigmoid respectively. The number of hidden layer nodes of Actor are 200,100 and 50, and the activation function are all tanh. Adam optimizer is used for learning the neural network parameters with a base learning rate of 0.001 and 0.003 for the Critic and Actor respectively. The discount factor $\gamma$ is set to 0.8, and the update parameter of target network $\tau$ is set to 0.1. $\aleph$ for action exploration is initialized 100, and the decreasing coefficient of learning is 0.99998.

**TABLE 3. The simulation results of the two SEMI-suspensions under the same conditions.**

| Road level velocity (km/h) | Percentage of Acceleration % | | Percentage of dynamic deflection% | | Percentage of vehicle body displacement% | |
|---|---|---|---|---|---|---|
| | SDC | DDPG | SDC | DDPG | SDC | DDPG |
| B,60 | 93.7 | 58.1 | 90.7 | 162.9 | 48.1 | 88.4 |
| B,120 | 98.0 | 66.3 | 98.6 | 140.8 | 45.1 | 95.9 |
| C,60 | 96.7 | 69.7 | 87.8 | 129.0 | 62.0 | 92.1 |
| C,120 | 99.4 | 69.6 | 100.8 | 137.9 | 37.9 | 107.2 |
| D,60 | 99.3 | 77.7 | 95.6 | 119.1 | 59.1 | 104.2 |
| D,120 | 99.0 | 86.4 | 97.8 | 126.9 | 43.4 | 106.0 |

## B. SIMULATION EXPERIMENT

### 1) SIMULATION EXPERIMENTS OF DIFFERENT ROAD LEVEL AND BODY VELOCITIES

In order to verify the control performance and robustness of the semi-active suspension system based on DDPG algorithm, simulation experiments have been carried out on the passive suspension system, the semi-active suspension system based on SDC and the semi-active suspension system based on DDPG algorithm under the same conditions. The simulation experiments were carried out on road level B, C and D at speeds of 60 km/h and 120 km/h respectively.

Figure 6 is part of the simulation experiment results about the vehicle body acceleration, suspension dynamic deflection and body displacement on level B at the speed of 60 km/h.

For convenience of comparison, the percentage of body vertical acceleration, suspension dynamic deflection and body displacement of the semi-active suspension system based on SDC and the semi-active suspension system based on DDPG algorithm to the relevant parameters of passive suspension system under the same conditions are given in table 3.

The simulation results show that compared with the semi-active suspension based on SDC, the semi-active suspension control method proposed in this paper has a certain loss in the dynamic deflection of the suspension for different road and different speeds, but it has a great improvement in the acceleration and the vehicle body displacement is not too bad. Therefore, the riding comfort of semi-active suspension based on DDPG algorithm is the best compared with the other two suspensions.

### 2) ROAD LEVEL SWITCHING TEST

In order to test the adaptability of the three suspension system proposed in this paper to different road levels, the simulation experiments were carried out when the vehicle body speed is 60 km/h and the road level is changed from C to D. Figure 7 is the simulation results.

The simulation results show that the average acceleration of the suspension system based on DDPG and system based on SDC is reduced by 24.7% and 2% respectively compared with the passive suspension system. And the

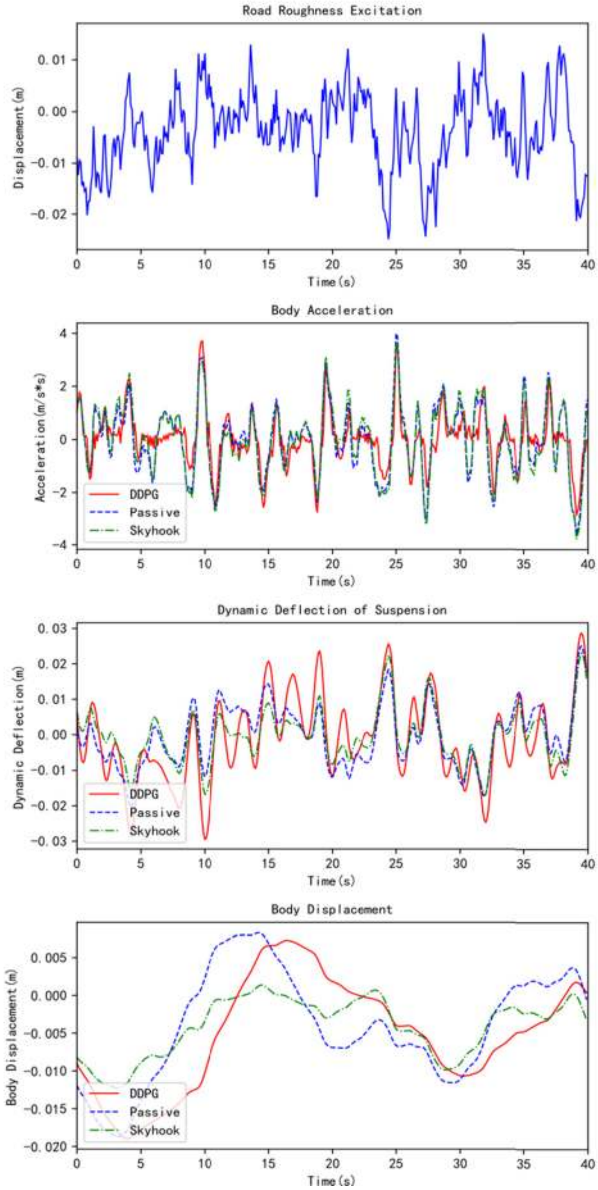



**FIGURE 6. Simulation results of the three suspensions at the same speed and road level.**

vehicle body displacement is reduced by 12.4% and 42.5%. The results show that the suspension system based on DDPG still has good performance when the road level is switched.

### 3) VEHICLE SPEED SWITCHING TEST

Figure 8 shows the simulation results under condition of switching the vehicle speed from 60 km/h to 120 km/h on road level C.

The simulation results show that the average acceleration of the suspension system based on DDPG and system based on SDC is reduced by 24.3% and 2.4% respectively compared with the passive suspension system. And the vehicle body displacement is reduced by 12.9% and 52.4%. The results

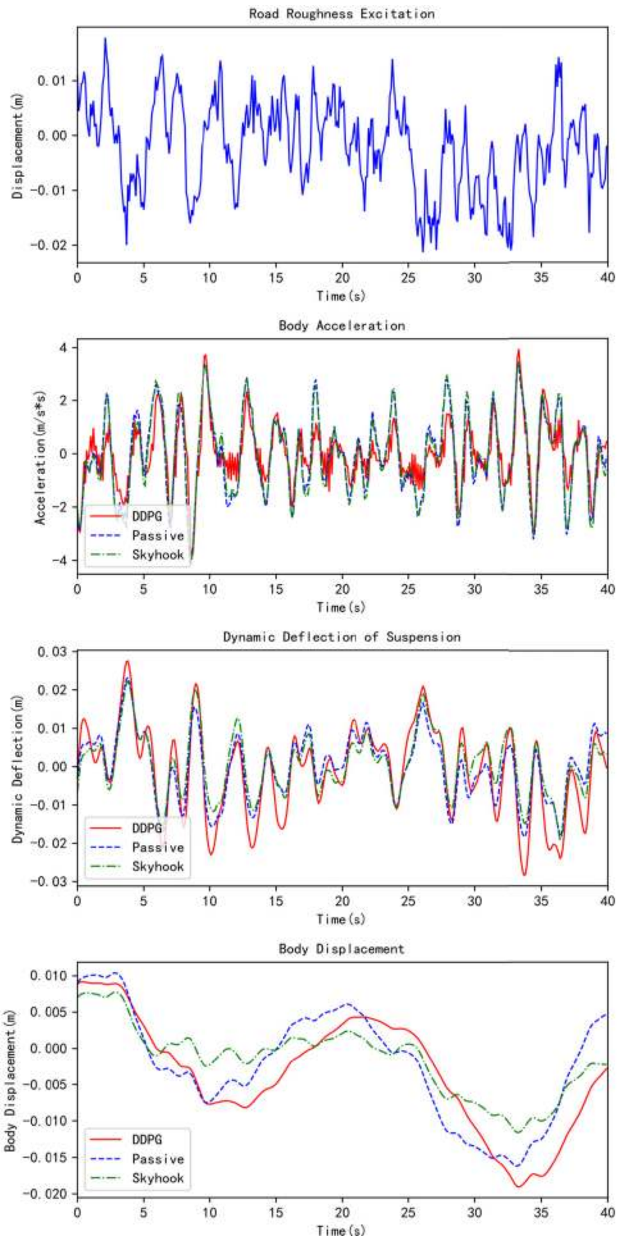



**FIGURE 7.** Simulation results of the three suspensions switching road level from C to D at the same speed.

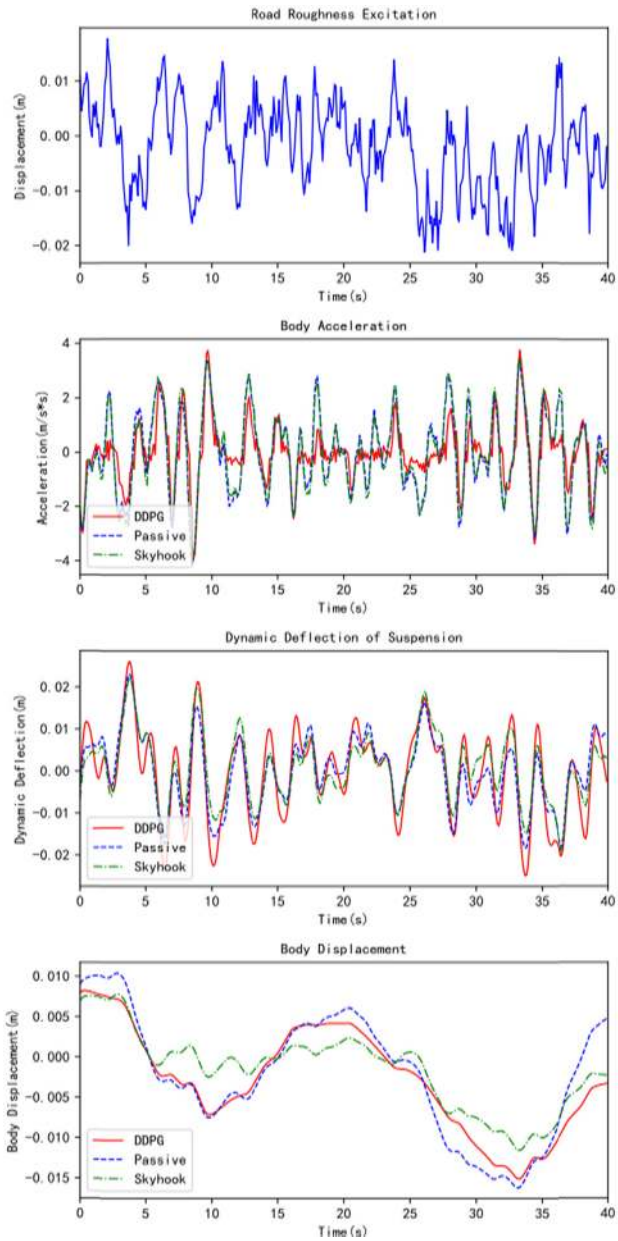


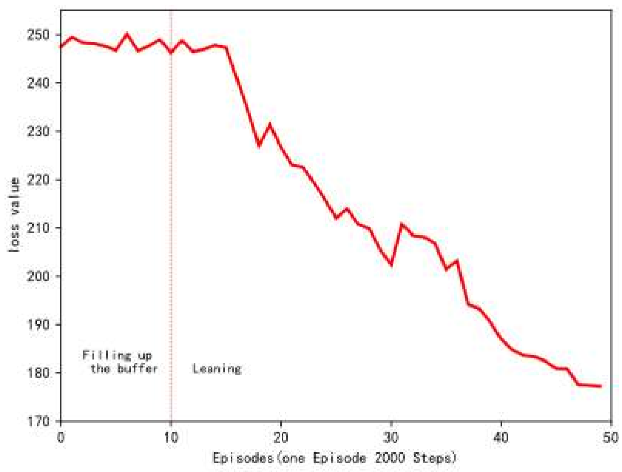

**FIGURE 8.** Simulation results of the three suspensions at the same road level with vehicle speed switching from 60km/h to 120km/h.

show that the suspension system based on DDPG still has good performance when the vehicle speed is switched. The simulation results show that the control method can adapt to the change of speed very well.

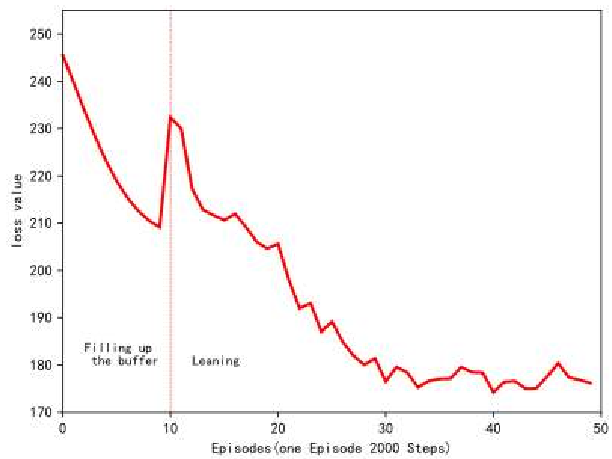
### C. COMPARISON OF DDPG LEARNING SPEED WITH EXPERIENCE SAMPLES

In inexperienced network learning, all the sample control quantities of filling buffer are obtained by random exploration. The initial sample of filling buffer with empirical learning is a group of samples with good control effect. The two groups of experiments set exactly the same learning parameters, and the learning convergence process is shown in figure 9.

Comparing the two groups of experiments, it is obvious that learning with experience converges faster at the beginning of learning. This shows that in the initial stage of learning, the samples that have achieved better performance have better learning value than those that are not based on systematic exploration. However, as the number of learning iterations increases, inexperienced learning can achieve the same good results, which shows that the final learning effect is determined by the learning structure, and has nothing to do with the empirical samples.




(a) Learning when no experience samples provided




(b) Learning with experience samples

**FIGURE 9.** **Comparison of learning efficiency.**

## VI. CONCLUSION

The active control system of semi-active suspension based on improved DDPG algorithm proposed in this paper can be well adapted to all kinds of road levels and speeds, and has good adaptability to different road switching and different vehicle speeds switching. The performance of this strategy is much better than that of passive suspension and the acceleration is much better than the SDC strategies too. The proposed learning approach with experienced samples can accelerate the speed of systematic training. The semi-active suspension system based on DDPG algorithm can also continuously learn according to external feedback information, so it can have better adaptability. However, the current results are only obtained through simulation studies. Whether the practical application has similar effects still needs our further research and exploration. In particular, the sensitivity parameter in establishing the vehicle-road-coupling equation for the vehicle dynamics model should be considered [26].

## REFERENCES

[1] J. L. Yao, S. Taheri, and S. M. Tian, "A novel semi-active suspension design based on decoupling skyhook control," *J. Vibroeng.*, vol. 16, no. 3, pp. 1318–1325, 2014.

[2] B. Lan and F. Yu, "Design and simulation analysis of LQG controller of active suspension," *Trans. Chin. Soc. Agricult. Mach.*, vol. 35, no. 1, pp. 13–18, 2004.

[3] Y. W. Yu, C. C. Zhou, and L. L. Zhao, "Design of LQG controller for vehicle active suspension system based on alternate iteration," *J. Shandong Univ., Eng. Sci.*, vol. 47, no. 4, pp. 50–58, 2017.

[4] R. W. Wang, X. P. Meng, D. H. Shi, "Fuzzy control of vehicle ISD semi-active suspension," *Trans. Chin. Soc. Agricult. Mach.*, vol. 44, no. 12, pp. 1–5, 2013.

[5] D. Y. Wang and H. Wang, "Control method of vehicle semi active suspensions based on variable universe fuzzy control," *China Mech. Eng.*, vol. 28, no. 3, pp. 366–372, 2017.

[6] J. L. Yao, W. K. Shi, J. Q. Zheng, "Development of a sliding mode controller for semi-active vehicle suspensions," *J. Vib. Control*, vol. 19, no. 8, pp. 1152–1160, 2013.

[7] H. P. Du and N. Zhang, "Constrained $H_\infty$ control of active suspension for a half-car model with a time delay in control," *Proc. Inst. Mech. Eng. D, J. Automobile Eng.*, vol. 222, no. 5, pp. 665–684, 2008.

[8] W. Sun, H. Gao, and O. Kaynak, "Finite frequency $H_\infty$ control for vehicle active suspension systems," *IEEE Trans. Control Syst. Technol.*, vol. 19, no. 2, pp. 416–422, Mar. 2011.

[9] W. Sun, Z. Zhao, and H. Gao, "Saturated adaptive robust control for active suspension systems," *IEEE Trans. Ind. Electron.*, vol. 60, no. 9, pp. 3889–3896, Sep. 2013.

[10] F. Zhao, S. S. Ge, M. Dong, F. Tu, and Y. Qin, "Adaptive neural network control for active suspension system with actuator saturation," *IET Control Theory Appl.*, vol. 10, no. 14, pp. 1696–1705, Sep. 2016.

[11] M. Milanese, C. Novara, and M. Canale, "Semi-active suspension control using 'fast' model-predictive techniques," *IEEE Trans. Control Syst. Technol.*, vol. 14, no. 6, pp. 1034–1046, Oct. 2006.

[12] D. Wang, D. Zhao, M. Gong, and B. Yang, "Research on robust model predictive control for electro-hydraulic servo active suspension systems," *IEEE Access*, vol. 6, pp. 3231–3240, 2018.

[13] S. H. Zhu, B. Z. Lv, and H. Wang, "Neural network control method of automotive semi-active air suspension," *J. Traffic Transp. Eng.*, vol. 6, no. 4, pp. 66–70, 2006.

[14] H. Li, Y. H. Feng, and L. Su, "Vehicle active suspension vibration control based on robust neural network," *Chin. J. Construct. Mach.*, vol. 15, no. 4, pp. 324–337, 2017.

[15] Y. Qin, J. Rath, C. Hu, C. Sentouh, and R. Wang, "Adaptive nonlinear active suspension control based on a robust road classifier with a modified super-twisting algorithm," *Nonlinear Dyn.* vol. 97, no. 4, pp. 2425–2442, Sep. 2019.

[16] S. K. Sharma, U. Saini, and A. Kumar, "Semi-active control to reduce lateral vibration of passenger rail vehicle using disturbance rejection and continuous state damper controllers," *J. Vib. Eng. Technol.*, vol. 7, no. 2, pp. 117–129, Apr. 2019.

[17] Q. Liu, J. Y. Zhai, and Z. C. Zhang, "A survey on deep reinforcement learning," *Chin. J. Comput.*, vol. 40, no. 1, pp. 1–28, 2017.

[18] D. Silver, A. Huang, and C. J. Maddison, "Mastering the game of Go with deep neural networks and tree search," *Nature*, vol. 529, no. 7587, pp. 484–489, 2016.

[19] P. Brezas, M. C. Smith, and W. Hoult, "A clipped-optimal control algorithm for semi-active vehicle suspensions: Theory and experimental evaluation," *Automatica*, vol. 53, pp. 188–194, Mar. 2015.

[20] Z. S. Yu, *Theory of Automobile*. Beijing, China: China Machine Press, pp. 234–235, 2018.

[21] Z. C. Wu, S. Z. Chen, and L. Yang, "Model of road roughness in time domain based on rational function," *Trans. Beijing Inst. Technol.*, vol. 29, no. 9, pp. 795–798, 2009.

[22] F. Lu and S. Z. Chen, "Modeling and simulation of road surface excitation on vehicle in time domain," *Automot. Eng.*, vol. 37, no. 5, pp. 549–553, 2015.

[23] V. Konda and N. Tsitsiklis, "Actor-critic algorithms," *Siam J. Control Optim.*, vol. 42, no. 4, pp. 1143–1166, 2003.

[24] T. P. Lillicrap, J. J. Hunt, and A. Pritzel, "Continuous control with deep reinforcement learning," in *Proc. Int. Conf. Learn. Represent.*, San Juan, Puerto Rico, 2016.

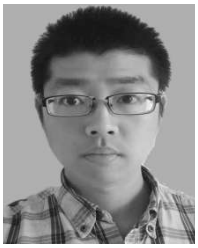
[25] M. Agostinacchio, D. Ciampa, and S. Olita, "The vibrations induced by surface irregularities in road pavements—A MATLAB approach," *Eur. Transp. Res. Rev.*, vol. 6, no. 3, pp. 267–275, Sep. 2014.

[26] R. C. Sharma and S. K. Sharma, "Sensitivity analysis of three-wheel vehicle's suspension parameters influencing ride behavior," *Noise Vib. Worldwide*, vol. 49, nos. 7–8, pp. 272–280, Jul. 2018.

**LIU MING** received the bachelor's and master's degrees from the Shandong University of Science and Technology, China, in 1995 and 1998, respectively. He is currently pursuing the Ph.D. degree with the School of Control Science and Engineering, Shandong University, China. He is currently an Associate Professor and the Vice Dean of the Department of Electrical Information, Shandong University of Science and Technology. His research interests include robotics and intelligent control.

**LI YIBIN** received the bachelor's degree from Tianjin University, China, in 1982, the master's degree from the Shandong University of Science and Technology, China, in 1988, and the Ph.D. degree from Tianjin University, in 2006. He is currently a Professor and the Associate Dean of the School of Control Science and Engineering, Shandong University, China. His research interests include robotics, mechatronics, and intelligent systems.

**RONG XUEWEN** received the bachelor's and master's degrees from the Shandong University of Science and Technology, China, in 1996 and 1999, respectively. He is currently pursuing the Ph.D. degree with the School of Control Science and Engineering, Shandong University, China. He is currently a Senior Engineer with the School of Control Science and Engineering, Shandong University. His research interests include robotics, mechatronics, and hydraulic servo driving technology.



**ZHANG SHUAISHUAI** received the bachelor's degree from the Shandong University of Technology, China, in 2009, and the master's and Ph.D. degrees from Shandong University, China, in 2012 and 2016, respectively. He is currently a Lecturer with the Department of Electrical Information, Shandong University of Science and Technology, China. His research interests include robotics technology and intelligent systems.

**YIN YANFANG** received the bachelor's and master's degrees from the Shandong University of Science and Technology, China, in 1999 and 2002, respectively, where she is currently pursuing the Ph.D. degree. She is currently a Lecturer with the Department of Electrical Information, Shandong University of Science and Technology. Her research interest includes machine learning.