Self-learning control of model uncertain active suspension systems with observer–critic structure

Zhijun Fu\(^1\), Peixin Yuan\(^1\), Fang Zhou\(^1\), Yaohua Guo\(^2\) and Pengyan Guo\(^3\)

Abstract
This paper presents a self-learning control algorithm for model uncertain suspension systems using single network adaptive critic (SNAC) approach. First, a differential neural network (DNN) observer in conjunction with the weight updating law is established to observe the uncertain dynamic. Then, the nominal optimal value function is approximated by a critic NN whose weight is updated by a novel design learning law driven by the filtered parameter error. The online self-learning control policy is thus derived by approximately solving the Hamilton–Jacobi–Bellman (HJB) equation based on SNAC technique. The Lyapunov approach is synthesized to ensure the convergent characteristics of the entire closed-loop system composed of the DNN observer and the self-learning control policy. Computer simulation of a quarter car suspension system is established to verify the effectiveness of the proposed approach. Simulation results illustrated that the designed method can ensure the good performance in terms with the road hold and ride quality. In addition, independent of model and online self-learning characteristics make it possible to design a high-performance vehicle active suspension controller.

Keywords
Observer, self-learning control, active suspension, single network adaptive critic

Introduction
A high-performance suspension system is the key to the pursuit of a smoother and safer car. Correspondingly, the suspension system is expected to have more intelligence to adapt to different road inputs. It needs to be pointed out that the suspension system is affected by unknown road input, resulting in unavoidable vibrations related to vehicle ride comfort. Therefore, when designing the controller of the active suspension system, state variable information related to the vertical dynamics of the vehicle should be provided.\(^5\) However, the implementation of suspension control system based on traditional observation theory brings great challenges because unknown road input is difficult to measure or observation when the vehicle is driving. Moreover, the uncertainty of parameters, such as sprung mass, also brings difficulties to the design of suspension observer and controller. Enhancing active vibration control performances as shown in refs. \(^3\) and \(^4\) is the main target of suspension control.

To deal with model uncertainty, robust approach,\(^5\) LMI based approach,\(^6\) sliding mode approach,\(^7,8,41\) and globally bounded Jacobian approach\(^9\) have been used to develop the observer. In ref. \(^10\), a robust decoupling observer considering disturbance is achieved, but the rank condition in terms with the output and input distribution matrix is difficult to meet. To avoid the requirement of rank constraints related to the system matrix, Sedighi et al.\(^11\) present a new development for the unknown input observers design, which guarantees the stability of the error closed-loop system and provides designers with a greater degree of freedom in design space. In ref. \(^12\), a controller and fault identification method based on the separation principle is proposed to realize the simultaneous decoupling identification and control. In ref. \(^13\), an adaptive extended state observer with lumped uncertainties and unmeasured states is proposed to estimate the unknown coefficient. However, most of the above observer designs rely on knowing the model information a priori, which poses a strict restriction to real application.

Neural network with good nonlinear approximation capability is very suitable to design the observer of model uncertain systems. The structure of a neutral observer mainly...
includes two parts, that is, a neural network to identify the unknown nonlinearity and a traditional Luenberger-like observer to estimate the state. In refs. 14 and 15, in the case of imposing strong strict positive real (SPR) conditions on the output error equation, a structure using two independent single layer neural networks is proposed to estimate the state of affine and non-affine SISO nonlinear systems. Furthermore, a nonlinear observer for the MIMO system using a single layer neural network was developed in refs. 16 and 17, where the SPR restrict condition has been weakened to a certain extent. Furthermore, a neural network observer with the modified backpropagation algorithm was proposed in ref. 18, where the SPR condition and any other strong restrictions are canceled. In particular, the differential neural network (DNN), incorporating the feedback design, provides a more efficient way to solve the state estimation problem of model uncertainty systems. In ref. 20, a DNN observer with sliding mode updating rule was reported and the relevant observational conditions have been removed. New passivity analysis of DNN proposed by Xiao et al.21 illustrates that the boundedness of external input can ensure the boundedness of the input–output signals for each block of closed-loop error dynamics, which therein avoids the requirement of persistency excitation (PE) condition.

Recent studies have shown that the design of a self-learning controller based on approximate dynamic programming (ADP) can avoid the requirement of model accuracy and achieve optimal control at the same time.22,23 As we all know, the curse of dimensionality of dynamic programming and offline learning characteristics make the existing optimal control strategies based on dynamic programming inefficient. ADP is inspired by biological systems, which can greatly improve the computing power for solving optimal control problems, but it will not bring obvious approximation errors. Since Werbos24 introduced a commonly used actor-critic (AC) scheme for self-learning control design, various modifications to ADP have been developed, such as action dependent heuristic dynamic programming (ADHDP),25 heuristic dynamic programming (HDP),26 dual heuristic programming (DHP),27 and Q-learning.28 The natural recursive characteristics of ADP are very suitable for solving discrete systems, and its related results cannot be directly used for reference in the continuous field. Ref. 29 proposes a self-learning control method for continuous nonlinear systems, where an identifier is used to identify unknown nonlinear dynamics, and then the online self-learning rate is obtained by approximately solving HJB through an evaluation network. However, it is limited to nonlinear affine system. As claimed in refs. 30 and 31, self-learning control method with the single network adaptive critic (SNAC) scheme has been proved to be an effective method to solve the HJB online and obtain the optimal control action.

It should be pointed that most of the above-mentioned self-learning control research are based on the premise that the state is fully known, which is a strong constraint condition, because in most cases the system state is not completely obtainable, such as the suspension system to be studied in this paper. The main contributions of this paper can be summarized as:

(1) It is the first time that the self-learning control of model uncertain suspension system using only the input and output of the system is investigated. Unlike the commonly used LQR controller design depending on the system model, this paper introduces the observer-based self-learning controller design without prior knowledge of the system dynamics, which makes the proposed method become more suitable for practical application.

(2) The investigated algorithm is implemented based on the observer–critic framework. First, a DNN observer with a reasonably designed weight online update law is used to identify unknown system dynamics based on the known system input/output information. Then, the self-learning control policy is derived with the help of SNAC method. With introducing two novel tuning laws for the DNN observer and the SNAC, the improved performance in terms of adaptability and robustness is realized compared with the general ADP-based self-learning method.

(3) In addition, the entire learning process of the proposed method was updated online, and the stability of the entire closed-loop system was guaranteed by properly designed composite Lyapunov method. The model free and self-learning characteristics of the designed self-learning control method make it possible to develop a good performance active suspension controller that can avoid the influence of uncertainty. The simulation results on a quarter car suspension model encountered unknown road input are presented to demonstrate the effectiveness of the designed self-learning control method.

The overall structure of this article is organized as follows. Section 2 presents the formulation of the question. Section 3 illustrates the DNN observer design process. Observer-based self-learning control method for a quarter car active suspension system is developed in the Section 4. Simulation results carried on a quarter car suspension system with time varying road input are given in Section 5. Finally, the conclusions are concluded in Section 6.

Problem formulation

As we all know, the quarter car suspension model as shown in Figure 1 is commonly used in the design of active suspension control systems.32 The state space equation considering the time varying parameters of a quarter car active suspension system can be expressed as

\[
\begin{align*}
\dot{x} &= (A + \Delta A)x + (B + \Delta B)u + Lz, \\
y &= Cx
\end{align*}
\]

(1)

where suspension state vector is denoted as \(x = [x_1, x_2, x_3, x_4]^T\), suspension deflection is denoted as \(x_1 = z_s - z_u\), sprung mass velocity is denoted as \(x_2 = \dot{z_s}\), tire deflection is denoted as \(x_3 = z_u - z_t\), unsprung mass velocity is expressed as \(x_4 = \dot{z_u}\), \(k_s\) is suspension stiffness, \(b_t\) represents suspension damper, active control input is express as \(u = f_d\), \(m_s\) is sprung mass which is equal to a quarter of the entire vehicle mass, \(m_u\) denotes unsprung mass, \(k_v\) denotes vertical stiffness of the tire, \(z_s\) is vertical displacement of sprung mass, \(z_u\) denotes vertical displacement of unsprung
mass, $z_r$ denotes vertical displacement of the unknown road, $\Delta A$ and $\Delta B$ represent the uncertainties caused by the uncertainty parameter $m_s$ and unknown input $z_r$,

$$
A = \begin{bmatrix}
0 & 1 & 0 & -1 \\
-k_x & -b_x & 0 & b_x \\
0 & 0 & 1 & 0 \\
-k_u & b_x & -k_u & b_x
\end{bmatrix}.
$$

The requirements of active suspension system controller design should meet the following three aspects:

1) The first main aspect is to ensure high-performance riding quality. This task aims to weaken the vibration force transmitted from the unsprung mass to the sprung mass through an advanced control approach, which is achieved by minimizing the sprung mass acceleration in the face of parameter uncertainties caused by sprung mass and time-varying road input.

2) In addition, the wheel should maintain firm and uninterrupted contact with the road surface, and the normal tire load change related to the vertical deflection of the tire should be small to ensure good road holding performance.

3) Finally, the suspension space constraint should be less than the maximum suspension deflection $z_{\text{max}}$, that is, $|z_s - z_u| \leq z_{\text{max}}$.

The goal of active suspension design is to find a control policy that can not only keep the system state stable, but also minimize the required performance indicator, such that

$$
V = \frac{1}{2} \int_0^t \left( x^T Q x + u^T R u \right) \, dt
$$

where the parameters $Q$ and $R$ represent different weights with respect to the state and control input, they are selected by the designer.

**Definition 1.** As claimed by Liews, if $u$ is continuous on $\Omega$, $u(0) = 0$, $u$ stabilizes (1) on $\Omega$, and $V(x_0)$ is finite, then the control input $u$ is called to be admissible with respect to (1) on compact set $\Omega \subseteq \mathbb{R}^n$ and is denoted by $u \in \mathcal{U}(\Omega)$.

With the help of knowledge in the field of optimal control, the Hamiltonian of (1) can be defined as

$$
H(x, u, V) = V_x^T [(A + \Delta A)x + (B + \Delta B)u] + Lz_r + r(x, u)
$$

where $V_x \triangleq \frac{\partial V(x)}{\partial x}$ denotes the partial derivative of the cost function $V(x)$ with respect to $x$.

The nominal optimal valued function $V^*(x)$ is expressed as

$$
V^*(x) = \min_{u \in \mathcal{U}(\Omega)} \int_0^\infty r(x(\tau), u(\tau)) \, d\tau
$$

Meanwhile, equation (4) should satisfy the following HJB equation

$$
0 = \min_{u \in \mathcal{U}(\Omega)} \left[ H(x, u, V^*_x) \right] = V_x^T [(A + \Delta A)x + (B + \Delta B)u] + Lz_r + r(x, u)
$$

where $V^*_x \triangleq \frac{\partial V^*(x)}{\partial x}$. Assume that the minimum of equation (5) uniquely exists, self-learning control policy $u^*$ can thus be obtained by solving $\partial H(x, u, V^*_x) / \partial u = 0$, such that

$$
u^* = -\frac{1}{2} \bar{R}^{-1} \left( \frac{\partial [(A + \Delta A)x + (B + \Delta B)u]}{\partial u} \right)^T V^*_x
$$

Then, the HJB equation (5) can be further expressed with respect to equation (6) as follows

$$
0 = \min_{u \in \mathcal{U}(\Omega)} \left[ H(x, u^*, V^*_x) \right] = V_x^T [(A + \Delta A)x + (B + \Delta B)u^*] + Lz_r + r(x, u^*)
$$

It can be seen that the nominal value function $V^*(x)$ should be known in advance to obtain the self-learning control policy $u^*$ in equation (6). However, the analytical solution of partial differential equation (7) cannot be obtained due to the unknown state, model uncertainty and unknown road input factors existed in the suspension system. This paper devotes to solve this problem by establishing an observer–critic framework via the single network adaptive critic method in the following two steps:

1) Based on the input/output information from the suspension system, an adaptive DNN observer with the weight updating law was established to observe the unknown state.

2) Based on the observations, a self-learning controller with the observer critic structure is developed via the single network adaptive critic approach.

**Remark 1.** The common LQR controller designs are often based on formula (2) where all parameters are available in advance. In addition, the corresponding feedback control law is obtained by solving the Riccati equation offline. The disadvantage of this method is that the control gain cannot be updated online according to the system uncertainties caused by $m_s$ and unknown road displacement $z_r$. This will inevitably lead to unsatisfactory control performance. Therefore, the design of active suspension systems is expected to introduce
a self-learning intelligent control method that can adapt to various driving conditions.

**The differential neural network observer**

Consider the following DNN to identify nonlinear system (1), such that

\[ \dot{x} = Ax + W\sigma([x^T, u^T])^T + \xi \]  \hspace{1cm} (8)

where \( W \in \mathbb{R}^{k \times d} \) represents the unknown nominal matrix, \( \sigma(\cdot) \) denotes the nonlinear activation function which is commonly chosen as a sigmoidal function \( \sigma(\cdot) = \frac{1}{1 + e^{-\alpha}} \) with the properly selected parameters \( \alpha, \beta, \gamma, \xi \) is regarded as modeling error.

**Assumption 1.** The nominal weight \( W \) and the approximate error vector \( \xi \) are bounded by

\[ \|W\|_F \leq \overline{W}, \|\xi\|_F \leq \overline{\xi} \] \hspace{1cm} (9)

Fact 1. Garces\(^{37}\) As we all know, the nonlinear activation function \( \sigma(\cdot) \) satisfies the generalized local Lipschitz conditions, such that

\[ \left\| \sigma([x^T, u^T]^T) - \sigma([\tilde{x}^T, u^T]^T) \right\| \leq \lambda_o\|x\|, (\lambda_o > 0) \] \hspace{1cm} (10)

Considering the following the neuro-observer

\[ \dot{\hat{x}} = A\hat{x} + \tilde{W}\sigma([\tilde{x}^T, u^T]^T) + K(Cx = C\hat{x}) \] \hspace{1cm} (11)

where \( \hat{x} \) denote the observation state, the observation gain \( K \) is obtained by solving the following stable equation

\[ (A - KC)^TP + P(A - KC) + 2\alpha P = -Q \] \hspace{1cm} (12)

where \( Q, P \) denote positive definite matrices and parameter \( \alpha \) will be explained in the subsequent proof analysis.

**The state observation error is expressed as**

\[ \dot{\hat{x}} = x - \hat{x} \] \hspace{1cm} (13)

From (8) and (11), the equation in terms with the observation error is derived as

\[ \dot{\hat{x}} = (A - KC)\hat{x} + W\sigma([\tilde{x}^T, u^T]^T) + \xi - \tilde{W}\sigma([\tilde{x}^T, u^T]^T) \]

\[ = (A - KC)\hat{x} + W\left\{ \sigma([\tilde{x}^T, u^T]^T) - \sigma([\tilde{x}^T, u^T]^T) \right\} \]

\[ + \xi + \tilde{W}\sigma([\tilde{x}^T, u^T]^T) \]

\[ \dot{\hat{y}} = C\hat{x} \] \hspace{1cm} (14)

where \( \tilde{W} = W - \tilde{W} \).

Once the architecture of the DNN observer is determined, the next step is to design appropriate update rules to implement online learning. There are generally the following two design ideas:

1) Based on commonly used learning rules, such as back propagation algorithm, the online learning rate of DNN observer is designed, and then suitable candidate of Lyapunov function is designed to prove the convergence characteristics of the system\(^{11,15}\)

2) In order to ensure the stability of the closed-loop system, the online learning rate is designed by defining the quadratic function which is related to the weight error and the observation error, and the time derivative of Lyapunov function is proved to be negative\(^{38,39}\).

The main drawback of the previous work in the first way is the approximate treatment of the dynamic backpropagation problem by using the gradient approximation, which inevitably leads to parameter overflow problems. Therefore, we choose the second way to develop the updating law for the propose DNN observer in this paper.

**Theorem 1.** The uncertain suspension model (1) is identified by the DNN observer model (11) with the following updating law

\[ \dot{\tilde{W}} = -\frac{\text{LPC}^T\overline{\xi}\sigma(\tilde{x}, u^T) - \lambda\overline{\xi}CP\|\tilde{W}\|}{\|C\|^2} \] \hspace{1cm} (15)

then the state estimation error and the DNN weight error is UUB.

That is, \( \dot{\hat{x}} \in L_\infty, \tilde{W} \in L_\infty \), where \( P \) is the solution of equation (12), \( L \) is positive definite matrix, \( \lambda \) is a positive constant.

**Proof.** The weight updating law in (15) is derived by properly selecting the following quadratic function related to the observation error and the weight error, such that

\[ L_t = \frac{1}{2}\tilde{x}^TP\tilde{x} + 2\alpha\{\tilde{W}^T L^{-1}\tilde{W}\} \] \hspace{1cm} (16)

The time derivative of \( L_t \) can be obtained by using the error dynamic equation (14), such that

\[ \dot{L}_t = \frac{1}{2}\tilde{x}^T\{P(A - KC) + (A - KC)^TP\}\tilde{x} \]

\[ + \tilde{x}^TP\tilde{W}\sigma([\tilde{x}^T, u^T]^T) - \sigma([\tilde{x}^T, u^T]^T) \]

\[ + \tilde{x}^TP\tilde{W}\sigma([\tilde{x}^T, u^T]^T) + \alpha\tilde{x}^TP\tilde{x} \] \hspace{1cm} (17)

Since the DNN weight is updated by (15) and satisfies the inequality \( \text{tr}(\tilde{W}^T L^{-1}\tilde{W}) < \|\tilde{W}\|^2 \) \( - \|\tilde{W}\|^2 \), then based on Assumption 1 and equation (16), \( L_t \) becomes

\[ \dot{L}_t \leq \frac{1}{2}\tilde{x}^T\{P(A - KC) + (A - KC)^TP\}\tilde{x} \]

\[ + \tilde{x}^TP\tilde{W}\sigma([\tilde{x}^T, u^T]^T) + \tilde{x}^TP\tilde{W}\sigma([\tilde{x}^T, u^T]^T) \]

\[ + \lambda\tilde{x}^TP\tilde{W} - \tilde{W}^T \]

\[ \leq \frac{1}{2}\tilde{x}^T\{P(A - KC) + (A - KC)^TP\}\tilde{x} \]

\[ + \lambda\tilde{x}^TP\tilde{W} + \tilde{x}^TP\tilde{W} \]

\[ + \lambda\tilde{x}^T\|P\|\|\tilde{W}\| - \|\tilde{W}\|^2 \] \hspace{1cm} (18)
Define \( \alpha = \lambda \bar{W} \) and select proper \( Q \) to satisfy the condition in (12), then we can get
\[
\dot{L}_i \leq -\frac{1}{2} \lambda \min \langle Q \rangle \Vert \tilde{x} \Vert^2 + \Vert P \Vert \| \tilde{x} \| \left( \tilde{z} + \mu^2 \right) \right) ^T (19)
\]
where \( \mu = \| \tilde{W} + \bar{\sigma} \| \).

The negative definiteness of \( \dot{L}_i \) can be guaranteed if
\[
\Vert \tilde{x} \Vert \geq \frac{2 \Vert P \Vert (\tilde{z} + \frac{\mu^2}{4\lambda})}{\lambda \min \langle Q \rangle} (20)
\]
Thus \( \dot{L}_i < 0 \) is guaranteed, which demonstrates the UUB of \( \| \tilde{x} \| \) by using the Lyapunov theory. The correctness of Theorem 1 is thus guaranteed.

Remark 2. Since we can select \( L, \lambda \) in (15) arbitrarily, the updating gains of DNN observer is not limited to the value of \( P \) which is obtained by solving the Riccati equation (12). Hence, one can select \( Q \) to make the state observation error in equation (20) as small as possible.

Observer-based self-learning controller design

This section is mainly devoted to the design of a self-learning control policy with the aid of the above-mentioned DNN observer using the single network adaptive critic approach.

The following single critic neural network is used to approximate the optimal value function, such that
\[
V^*(\tilde{x}) = W_c^T \psi(\tilde{x}) + \xi_c (21)
\]
where \( \psi \in \mathbb{R}^1 \) is the nominal weight vector, the activation function \( \psi \) is selected as the sigmoid nonlinear function, \( f \) is the number of neurons and \( \xi_c \) denotes the approximation error.

The derivative of the optimal value function (21) in terms of \( \tilde{x} \) is
\[
\bar{V}_c = \nabla \psi^T(\tilde{x}) \bar{W}_c + \nabla \xi_c (22)
\]
where \( \nabla \psi(\tilde{x}) = \frac{\partial \psi(\tilde{x})}{\partial \tilde{x}} \) and \( \nabla \xi_c = \frac{\partial \xi_c}{\partial \tilde{x}} \) represent the partial derivative of \( \psi(\tilde{x}) \) and \( \xi_c \) in terms with \( \tilde{x} \), respectively.

By substituting (22) into (7), the self-learning control policy is derived as
\[
u^* = -\frac{1}{2} \mathbb{R}^1 \left( \frac{\partial f}{\partial u} \left( A + \Delta A \right) \tilde{x} + (B + \Delta B)u + Lz \right) \right) ^T \nabla \psi^T(\tilde{x}) \bar{W}_c + \nabla \xi_c (23)
\]
Then, the approximate self-learning control is represented as
\[
u = -\frac{1}{2} \mathbb{R}^1 \left( \frac{\partial f}{\partial u} \left( A + \Delta A \right) \tilde{x} + (B + \Delta B)u + Lz \right) \right) ^T \nabla \psi^T(\tilde{x}) \bar{W}_c + \nabla \xi_c (24)
\]
\[
\nabla \psi^T(\tilde{x}) \bar{W}_c + \nabla \xi_c (25)
\]
Next, after substituting equation (22) into equation (6), the HJB equation can be further obtained as follows
\[
0 = \left( \nabla \psi^T(\tilde{x}) \bar{W}_c + \nabla \xi_c \right) \left( A + \Delta A \right) \tilde{x} + (B + \Delta B)u + Lz + \bar{r}(\tilde{x}, u) (26)
\]
where \( \bar{r}(\tilde{x}, u) \) is designed as
\[
\bar{r}(\tilde{x}, u) \equiv -\lambda \min \langle Q \rangle \| \tilde{x} \| (\tilde{z} + \mu^2) \right) \right) ^T \nabla \psi^T(\tilde{x}) \bar{W}_c + \nabla \xi_c (27)
\]
Remark 3. Our previous study\(^{40}\) show that, compared with the traditional observer design driven by observation error, the adaptive learning law including parameter error information can make the closed-loop system converge as soon as possible. Inspired by this result, the following analysis will be devoted to develop a novel learning law for critic neural networks driven by parameter errors, rather than the existed least square\(^{33}\) or gradient methods.\(^{41}\)

The auxiliary variables \( X \) and \( Y \) are defined, \( Y_f \) as
\[
\left\{ \begin{array}{l}
\eta Y_f + Y_f = Y, Y_f(0) = 0 \\
\eta X_f + X_f = X, X_f(0) = 0
\end{array} \right. (28)
\]
Further, another two auxiliary regression variables \( G \) and \( H \) are designed as
\[
\left\{ \begin{array}{l}
\bar{G}(t) = -\eta G(t) + X_f \int_0^t \frac{Y - Y_f}{\eta} dr \\
\bar{H}(t) = -\eta H(t) + X_f \int_0^t \frac{Y - Y_f}{\eta} dr
\end{array} \right. (29)
\]
By solving the differential equation shown in (28), the following expression can be obtained
\[
\left\{ \begin{array}{l}
G(t) = \int_0^t e^{-\eta(t-t')} X(r) X_f(t') dr \\
H(t) = \int_0^t e^{-\eta(t-t')} X(r) \left( \frac{Y - Y_f}{\eta} \right) dr
\end{array} \right. (30)
\]
In the end, the learning rule of \( \bar{W}_c \) is designed as
\[
\dot{\bar{W}} = -\mu \bar{J} (29)
\]
where \( \mu \) represents the adaptive updating gain and the filtered vector \( \bar{J} \) is defined as \( \bar{J} = G(t) \bar{W}_c + H(t) \).

Theorem 2. A control mechanism composed of a DNN observer (11) with an update law (15) and a self-learning control (23) with a learning law (30) is used to control the
system (1), then all the signals $\dot{x}, \dot{W}, \ddot{W}$ implied in the entire learning process consisting of the DNN observer, the SNAC and the self-learning control policy are UUB.

**Proof.** Let us design the following composite Lyapunov function

$$L = L_1 + L_c$$

(31)

where $L_1$ is defined as in equation (16) and $L_c$ is defined as

$$L_c = \frac{1}{2}\dot{W}_c^T\mu^{-1}\dot{W}_c$$

(32)

From (29), we have

$$J = G(t)W_c + H(t) = -G(t)\dot{W}_c + \dot{z}_f$$

(33)

where $z_f = \int_{0}^{t} e^{-\mu(t-r)}X_t\dot{z}_{HJ}\,dr$ is assumed to be bounded, that is, $\|z_f\| \leq \bar{z}_f$.

As claimed by Na,42 if the $X$ satisfies the persistently excited (PE) condition, then the matrix $G(t)$ defined in (29) is positive, that is, $\lambda_{\min}(E) > \sigma > 0$. Therefore, according to the equation $\dot{W}_c = -W_c$, the derivative of $L_1$ is expressed as

$$\dot{L}_1 = \dot{W}_c^T\mu^{-1}\dot{W}_c - E(t)\dot{W}_c^T\dot{W}_c + \dot{W}_c\dot{x}_f$$

(34)

Based on the well-known Cauchy’s inequality, that is, $ab \leq a^2\delta/2 + b^2/2\delta$ with $\delta > 0$, (34) is rewritten as

$$\dot{L}_1 \leq -\left(\sigma - \frac{1}{2\delta}\right)\|\dot{W}_c\|^2 + \frac{\delta\dot{z}_f^2}{2}$$

(35)

Then, from (8) and (32), we have

$$L_c = 2\Gamma_1^T\left(Ax + W\sigma(x^T, u^T)^T + \dot{z}_f\right) + \kappa V(x)$$

$$\leq \Gamma(2\|A\| + 2)\|x\|^2 + \Gamma\left|W\sigma(x^T, u^T)^T\right|^2$$

$$\leq \Gamma\left|z_f^2 - \Gamma\lambda_{\min}(Q)\|C\|^2\|x\|^2 - \Gamma\lambda_{\min}(R)\|u\|^2\right|^2$$

$$\leq -\left(\Gamma\lambda_{\min}(Q)\|C\|^2 - \Gamma(2\|A\| + 2)\|x\|^2\right) + \Gamma(\lambda_{\min}(Q)\|C\|^2 - \Gamma(2\|A\| + 2)\|x\|^2)$$

$$\leq -\Gamma\lambda_{\min}(Q)\|C\|^2 + \Gamma\lambda_{\min}(R)\|u\|^2 + \Gamma W^2\sigma^2$$

(36)

Then the derivative of $\dot{L}$ can be obtained from (16), (35), and (36), such that

$$\dot{L} \leq -\left(\sigma - \frac{1}{\delta}\right)\|\dot{W}_c\|^2 - \Gamma\lambda_{\min}(Q)\|C\|^2$$

$$-\Gamma\left(2\|A\| + 2\right)\|x\|^2 - \Gamma\lambda_{\min}(R)\|u\|^2$$

$$-\frac{1}{2}\lambda_{\min}(Q)|x|^2 + \|P\|\left(\dot{z}_f^2 + \left|\dot{x}\right|^2 + \frac{\|\mu\|^2}{4\lambda} + \sigma\right)$$

(37)

where $\sigma = \Gamma W^2\sigma^2 + e^{-1}\Gamma W^2\sigma^2 + \frac{\sigma}{2}$ is a positive constant.

With the assumption that the following conditions hold

$$\kappa > \Gamma(2\|A\| + 2)\lambda_{\min}(Q)\|C\|^2$$

$$\|\dot{W}_c\| > \sqrt{\left(\mu - 1\right)\delta}$$

(38)

Then, $\dot{L} < 0$ is proved. Hence, we concluded that UUB stability in terms with the DNN weight error $\dot{W}$ and critic NN weight error $\dot{W}_c$ implied in the whole closed-loop systems are guaranteed. Thus, Theorem 2 is proved.

Based on the above analysis, the mechanism of the DNN observer-based self-learning control approach for model uncertain suspension systems is shown in Figure 2.

**Remark 4.** A self-learning control method is proposed for uncertain suspension system. The significant advantage of the proposed method is that it gets rid of the limitations of the model, which is achieved by employing an DNN observer to identify the unknown dynamic and system state, where the weight updating law obtained by properly designed Lyapunov function instead of the commonly used backpropagation algorithm.43 Moreover, all the signals implied in the entire learning process are UUB.

Based on the above analysis, the flowchart of the proposed control algorithm is depicted in Figure 3 and can be summarized as follows.

1. Select the proper initial values of active functions $\sigma(\cdot)$, observation gain $K$ in equation (11) and updating gains $\lambda, \mu$ in equation (15) for the observer. $\sigma(\cdot)$ is usually selected as the sigmoidal function $\sigma(\cdot) = a/(1 + e^{-b\cdot}) - c$ where $a, b, c$ are the designed constants. $W$ and $\dot{W}_c$ are tuned online according equations (15) and (30). Hence, there is no need to select the initial values of $W$ and $\dot{W}_c$. Meanwhile, select the proper function $\psi(\cdot)$ in equation (21) and the updating gain $\mu$ in equation (30) for the critic SNAC. $\psi(\cdot)$ is usually selected as a smooth function consisting with the different combination between the selected suspension states.

2. Then, the inputs/outputs data of the suspension system is used to train the neural network including the DNN observer in equation (11) and critic NN in equation (21).

3. Finally, self-learning control law expressed in equation (24) is obtained based on the first two steps.

**Simulation and analysis**

In this section, a car active suspension model platform presented in Section 2 is numerically simulated to verify the effectiveness of the DNN observer designed in Section 3 and the self-learning controller developed in Section 4. Table 1 shows the relevant suspension parameters used in the simulation process.

In order to verify the robustness of the proposed method, we adopt the following random road input model as following
As claimed by Na, if the derivative of \( L_c \) is defined, that is,

\[
\dot{L}_c = \frac{d}{dt} L_c = \frac{d}{dt} \left( \sigma \right) = \sigma \dot{\sigma} + \frac{1}{2} \dot{\sigma}^2 + w_c \sigma^2 + \sigma^3 \quad (34)
\]

Then the derivative of \( L_c \) is given by

\[
\dot{L}_c = \sigma \dot{\sigma} + \frac{1}{2} \dot{\sigma}^2 + w_c \sigma^2 + \sigma^3 \quad (35)
\]

where the weight vector of the SNAC is represented by \( W_c \).

The flowchart of the proposed control algorithm is depicted in Figure 3 and can be summarized as follows.

1. Start
2. Initialization
   - \( \sigma, K \): Active function and observation gain for DNN observer in (11)
   - \( \psi \): Active function for Critic NN in (21)
   - \( L, \lambda \): Constants as in (15)
   - \( \mu \): Constant in (30)
3. Input/output system data
4. Observation model (11) with updating law (15)
5. SNAC model with Learning law (30)
6. Self-learning control law (24)
7. End

The simulations, such that

\[
V(x) = W_c x_1^2 + W_c x_2 x_3 + W_c x_2^2 x_3 + W_c x_3^4 + W_c x_4^4
\]

where the weight vector of the SNAC is represented by \( W_c = [W_{c1}, W_{c2}, W_{c3}, W_{c4}, W_{c5}, W_{c6}, W_{c7}, W_{c8}, W_{c9}, W_{c10}] \).

Random uncertain road input considering pavement roughness as showed in (39) is performed in this paper to compare the effectiveness of the proposed control method with the LQR control method. Moreover, time varying parameter of the longitudinal velocity is also considered in the simulations, such that

Case A. Simulations results with the road input (39) and velocity 20 km/h are illustrated in Figures 4 to 7. The simulation results indicated that the proposed self-learning control approach is more effective than the LQR control suspension, the ride quality has been improved observably with respect to the suspension deflection, sprung mass acceleration, tire deflection and unsprung mass velocity, which means the suspension system can be brought back to the stable state as fast as possible when encountering unknown road input. The suspension performances in terms with the road hold quality is thus greatly improved.

Case B: Simulations results with the road input (39) and velocity \( v = 60 \) km/h are presented in Figures 8–11. One can easily find that the proposed control method still has better performance with smaller fluctuations when encountering different longitudinal speeds compared with the commonly used LQR control method. These facts further verify the strong robustness and self-adaptive properties of the proposed control method. It should be mentioned that there are two main reasons for the performance improvements with the proposed control method in the simulation results above. First, the control law of the LQR method is based on the accurate suspension model, the control accuracy cannot be guaranteed when subjecting to time varying parameters. However, the proposed control method is not limited to a mathematical model, but based on the inputs/outputs data of suspension system. Second, the feedback control law of the proposed method can be updated online with the time varying parameters like longitudinal speed and random road input, whereas for the LQR method and other model-based methods, the feedback control law is fixed in advance. Therefore, it could be concluded that self-adaptive property of the proposed method provides a more effective solution for
the suspension controller design and can greatly enhance the vehicle riding performance.

To show the control performance of the proposed method, the performance index—Root Mean Square (RMS) for the states error has been adopted for the purpose of comparison.

\[ V(x) = W_{c1}x_1^2 + W_{c2}x_1x_2 + W_{c3}x_1x_3 + \ldots + W_{c10}x_3x_4 \]

where \( x \) is number of the simulation steps, \( W_c = [W_{c1}, W_{c2}, W_{c3}, W_{c4}, W_{c5}, W_{c6}, W_{c7}, W_{c8}, W_{c9}, W_{c10}]^T \) is the difference between the state variables with control and without control at \( i^{th} \) step.

One can also notice from Table 2 that the RMS values of all state variables for the proposed method are smaller than the commonly used LQR method, which further demonstrates the improved performance of the proposed method.

The above simulation results using the proposed DNN observer and self-learning controller indicated that all the evaluation indexes of suspension systems have been improved compared with the commonly used LQR control.
method. The proposed observer-based self-learning mechanism can realize online self-renewal according to the unknown road displacement without the complete suspension model. It therefore concluded that self-learning and model independent characteristics of the developed self-learning control approach opens a new idea for the design of model uncertain suspension control system and can apparently improve the car road hold and ride quality. Therefore, the aim of pursuing a high-performance suspension system is achieved.

Conclusions

This paper presents a new way to realize simultaneous online state observation and control for model uncertain suspension systems. This article has achieved the following main innovations. First, the proposed self-learning control method does not require the complete information of suspension system, this feature makes it become easier to be applied in practical system. Second, the self-learning control policy can be updated online with the unknown road input, which can greatly enhance the suspension performances in terms with the road hold and ride quality. Finally, simulation results performed on a quarter car suspension system considering unknown road input are presented. The suspension responses with respect to suspension deflection, spring mass acceleration, tire deflection and velocity of unsprung mass are presented to validate the effect of the proposed approach. The future work will focus on developing a more practical intelligent controller for automotive suspension systems considering actual state constraints and actuator saturation limits.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work is supported by the National Natural Science Foundation of China (Grant No.62073298), Key Scientific and Technological Project of Henan Province (Grant No. 212102310454 and No. 212102210237) and ZZULI Doctoral Fund for Scientific Research (Grant No. JDG20190099).

ORCID iD

Zhijun Fu https://orcid.org/0000-0002-9147-3676

References

1. Chen T, Cao D, Yuan J, et al. Observer-based adaptive active network backstepping sliding mode control for switched fractional order uncertain nonlinear systems with unmeasured state. Meas Control 2021; 54(7–8): 1–14. DOI: 10.1177/0036602111021107
2. Kim G, Kang S, Kim J, et al. Simultaneous estimation of state and unknown road roughness input for vehicle suspension control system based on discrete Kalman filter. Proc Imeche D: J Automobile Eng 2020; 234(6): 1610–1622.
3. Ebrahimi F, Nopour R and Dabbagh A. Smart laminates with an auxetic ply rested on visco-Pasternak medium: Active control of the system's oscillation. Engineering with Computers, 2021, DOI: 10.1007/s00366-021-01533-1
4. Zhang Y, Wang Z, Tazeddinova D, et al. Enhancing active vibration control performances in a smart rotary sandwich thick nanostructure conveying viscous flow by a PD controller Waves in Random and Complex Media, 2021, DOI: 10.1080/17455030.2021.1948627
5. Alif A, Darouach M and Boutayeb M. Design of robust reduced-order unknown-input filter for a class of uncertain linear neutral system. IEEE Trans Automatic Control 2010; 55(1): 6–19.
6. Zemouche A and Boutayeb M. On LMI conditions to design observers for Lipschitz nonlinear systems. Automatica 2013; 49: 585–591.
7. Kalsi K, Lian J, Hue S, et al. Sliding-mode observers for systems with unknown inputs: A high-gain approach. Automatica 2010; 46(2): 347–353.
8. Chu Z, Chen Y, Zhu D, et al. Observer-based adaptive neural sliding mode trajectory tracking control for remotely operated vehicles with thruster constraints. *Trans Inst Meas Control* 2021; 43(13): 2969–2971.
9. Phanomchonng G and Rajamani R. New rollover index for the detection of tripped and untripped rollovers. *IEEE Trans Ind Elect* 2013; 60(10): 4726–4736.
10. Rocha-Cozatl E and Moreno A. Dissipative design of unknown input observers for systems with sector nonlinearities. *Int J Robust Nonlinear Control* 2011; 21: 1623–1644.
11. Liu Z. Adaptive extended state observer-based heading control for spacecraft attitude control systems. *Chin J Electr* 2018; 27(1): 102–109.
12. Abdollahi F, Talebi H and Patel R. A stable neural network-based observer with application to flexible-joint manipulators. *IEEE Trans Neural Networks* 2006; 17(1): 118–129.
13. Talebi H, Abdollahi F, Patel R, et al. Neural network-based state estimation of nonlinear systems application to fault detection and isolation. *Springer Science Business Media, LLC*, 2010.
14. Xu J, Na J, Liu Z, et al. Adaptive state observer for uncertain MIMO systems with uncertain dynamics. Chinese Control and Decision Conference 2010; 1017–1022.
15. Khoygani R and Ghasemi M. Neural estimation using a stable discrete-time MLP observer for a class of discrete-time uncertain IMIMO nonlinear systems. *Nonlinear Dyn* 2016; 84: 2517–2533.
16. Wen X, Wang J and Li X. A feed-forward wavelet neural network adaptive observer-based fault detection technique for spacecraft attitude control systems. *Chin J Electr* 2018; 27(1): 102–109.
17. Poznyak A, Sanchez E and Yu W. Differential neural networks for robust nonlinear control, identification, state estimation and trajectory tracking. *Singapore*: World Scientific Publishing Co. Pte. Ltd., 2001.
18. Yu W, Marino M and Li X. Observer-based neural identifier. *IEEE Proc Control Theor Appl* 2000; 147(2): 145–152.
19. Xiao J, Zeng Z, Shen W, et al. Passivity analysis of delayed neural networks with discontinuous activations via differential inclusions. *Nonlinear Dyn* 2013; 74: 213–225.
20. Lewis L and Liu D. Approximate dynamic programming and reinforcement learning for feedback control. *Hoboken, NJ: Wiley*, 2013.
21. Khodamipour G, Khorashadizadeh S and Farshad M. Observer-based adaptive control of robot manipulators using reinforcement learning and the Fourier series expansion. *Trans Inst Meas Control* 2021; 43(10): 2307–2320.
22. Werbos P. Approximate dynamic programming for real-time control and neural modeling. In: *Handbook of Intelligent Control, Neural, Fuzzy, and Adaptive Approaches*. New York: Van Nostrand Reinhold, 1992.
23. Zhu L, Modares H, Peen G, et al. Adaptive suboptimal output-feedback control for linear systems using integral reinforcement learning. *IEEE Trans Contr Syst Technol* 2015; 23(1): 264–273.
24. Miller T, Sutton R and Werbos P. *Neural Networks for Control*. Cambridge, MA: MIT Press, 1990.
25. Fairbank M, Alonso E and Prokhorov D. Simple and fast calculation of the second-order gradients for globalized heuristic dynamic programming in neural networks. *IEEE Trans Neural Netw* 2012; 23(10): 1671–1676.
26. Wei Q, Liu D and Shi G. A novel dual iterative Q-learning method for optimal battery management in smart residential environments. *IEEE Trans Ind Elect* 2015; 62(4): 2509–2518.
27. Bhasin S, Kamalapurkar R, Johnson M, et al. A novel actor-critic-identifier architecture for approximate optimal control of uncertain nonlinear systems. *Automatica* 2013; 49(1): 82–92.
28. Lv Y, Na J, Yang Q, et al. Online adaptive optimal control for continuous-time nonlinear systems with completely unknown dynamics. *Int J Control* 2016; 89(1): 99–112.
29. Yang X, Liu D and Wang D. Reinforcement learning for adaptive optimal control of uncertain nonlinear systems with input constraints. *Int J Control* 2014; 87(3): 553–566.
30. Rajamani R. *Vehicle dynamics and control*. Springer, 2012.
31. Khalaf A and Lewis L. Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach. *Automatica* 2005; 41(5): 779–791.
32. Brezas P and Smith M. Linear quadratic optimal risk-sensitive control for vehicle active suspensions. *IEEE Trans Control Syst Technol* 2014; 22(2): 543–556.
33. Kumar M and Vijayarangan S. Design of LQR controller for active suspension system. *Indian J Eng Mater Sci* 2006; 13: 173–179.
34. Nguyen L, Park S, Turnip A, et al. Application of LQR control theory to the design of modified skyhook control gains for semi-active suspension systems. *ICCAS-SICE*, 2009, pp. 4698–4703.
35. Garces F, Becerra V, Kambhampati C, et al. *Strategies for feedback linearization: A dynamic neural network approach*. New York: Springer-Verlag, 2003.
36. Fu Z, Xie W and Na J. Robust adaptive nonlinear observer design via multi-time scales neural network. *Neurocomputing* 2016; 190: 217–225.
37. Tobias B and Karl K. Neural network based nonlinear observers. *Syst Control Lett* 2021; 148: 1–10.
38. Fu Z, Xie W, Rakheja S, et al. Adaptive optimal control of nonlinear systems with different time scales. *Neurocomputing* 2017; 238: 179–190.
39. Zhang X, Zhang H, Sun Q, et al. Adaptive dynamic programming-based optimal control of unknown nonaffine nonlinear discrete-time systems with proof of convergence. *Neurocomputing* 2012; 91: 48–55.
40. Na J, Xing Y and Costa-Castelló R. Adaptive estimation of time-varying parameters with application to roto-magnet plant. *IEEE Trans Syst Man, Cybernetics: Syst* 2021; 51(2): 731–741.
41. Kosko B. Bidirectional associative memories: unsupervised hebbian learning to bidirectional backpropagation. *IEEE Trans Syst Man, Cybernetics: Syst* 2021; 51(1): 103–115.
42. Dhananjay K and Kumar S. Modeling and simulation of quarter car semi active suspension system using LQR controller. *Adv Intell Syst Comput* 2014; 1: 441–448.