IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received March 24, 2022, accepted March 30, 2022, date of publication April 8, 2022, date of current version April 18, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3165937

Robust H_∞ Observer-Based Reference Tracking Control Design of Nonlinear Stochastic Systems: HJIE-Embedded Deep Learning Approach

BOR-SEN CHEN[1,2], (Life Fellow, IEEE), AND PO-HSUN WU[1]
[1]Department of Electrical Engineering, National Tsing Hua University, Hsinchu 30013, Taiwan
[2]Department of Electrical Engineering, Yuan Ze University, Chung-Li, Taoyuan 32003, Taiwan

Corresponding author: Bor-Sen Chen (bschen@ee.nthu.edu.tw)

This work was supported by the Ministry of Science and Technology of Taiwan under Grant MOST 108-2221-E-007-099-MY3.

ABSTRACT The robust H_∞ observer-based reference tracking control design of nonlinear stochastic systems with external disturbance and measurement noise is always a very complicated and difficult problem in the control field. It needs to solve a very difficult control-observer-coupled Hamilton Jacobi Isaacs equation (HJIE) for nonlinear observer and controller in the design procedure. At present, there exists no analytic and numerical way for solving this control-observer-coupled HJIE. A novel HJIE-embedded deep learning approach is proposed as a co-design of deep learning algorithm and H_∞ observer-based tracking control scheme to directly solve the nonlinear partial differential control-observer-coupled HJIE of H_∞ observer-based reference tracking control design problem of nonlinear stochastic systems. In the off-line training phase, state estimation error and tracking error are inputed to HJIE-embedded deep neural network (DNN) to output the solution of HJIE. If not, the learning error of HJIE is fedback to train DNN to solve HJIE for H_∞ tracking control law, observer gain as well as the worst-case external disturbance and measurement noise, which will be sent back to nonlinear stochastic system model to replace the external disturbance and measurement noise and estimation error signal for next step training. The proposed DNN-embedded H_∞ observer-based reference tracking scheme can achieve the theoretical H_∞ observer-based reference tracking control strategy as the deep learning algorithm converges. If free of external disturbance and measurement noise, the proposed DNN-based H_∞ observer-based reference tracking control scheme can approach to the stochastically asymptotical state estimation and reference tracking simultaneously. Finally, a design example of H_∞ observer-based reference tracking control for quadrotor UAV system with external disturbance and output measurement noise is provided to illustrate the design procedure and to validate the state estimation and reference tracking performance simultaneously of the proposed HJIE-embedded H_∞ DNN-based observer-based reference tracking control scheme of nonlinear stochastic systems.

INDEX TERMS Nonlinear stochastic system, nonlinear H_∞ observer-based output feedback reference tracking control, Hamilton-Jacobi Isaacs equation (HJIE), HJIE-embedded DNN-based observer-based control scheme, Adam learning algorithm, quadrotor UAV.

I. INTRODUTCTION

Deep neural network (DNN) was inspired by biological neural systems as an information processing model [1]. Through learning from big data, DNN enables us to perform tasks for a large variety of applications. Recently, many powerful big data-driven methods have been developed based on DNN and exceptionally applied to speech recognition [2], [3], translation of languages [3], [6] and image classification [4], [5], etc. These works like speech recognition, language translation and image classification are usually simple to perform by human brain but are still difficult for man-made machines. For the above applications, we need to train DNN as best as possible in order to fit the above input/output data pairs [7], [8]. Therefore, the big data-driven method has been employed for the current training methodology of

The associate editor coordinating the review of this manuscript and approving it for publication was Yilun Shang.

deep learning methods of DNN [9]. Once it has been trained, DNN can respond to those never-observed data to make an optimal recognition, translation or classification according to its past trained knowledge [8], [9].

In the last few years, due to the development of hardware to process a very large amount of data, DNN-based learning methods have been significantly improved with wide applications, which require a very large amount of empirical data to train for a specific optimal performance [7], [9]. At present, the traditional DNN-based deep learning methodologies always employ big data-driven approaches since they need a very large amount of empirical data to train DNN for the system knowledge and behavior [1], [9]. The shortage of these traditional big data-driven deep learning methodologies always neglects the traditional system modelings and a large amount of well-developed theoretical results. For example, unlike speech recognition and image classification, the stochastic nonlinear system models with external disturbance and measurement noise have been constructed and the corresponding theoretical robust H_∞ state estimation and observer-based output feedback control results of nonlinear stochastic systems have been well developed. But an analytic or efficient method is still lack to solve them under intense research in the field of system control for several decades [10]–[14]. Therefore, it is more appealing to apply the deep learning scheme to solve these complicated nonlinear H_∞ stochastic observer-based control design problems other than the traditional big data-driven deep learning methods for speech recognition and image classification problem. These theoretical results of robust H_∞ state estimation and observer-based output feedback control design of nonlinear stochastic system under external disturbance and measurement noise could be considered as expert knowledge in the deep learning approaches, which can save much training data and time in the DNN-based robust H_∞ observer-based control design of nonlinear stochastic system with external disturbance and output measurement noise.

In the last decades, the robust H_∞ control strategies have been developed and widely applied to efficiently attenuate the effect of external disturbance on the stabilization performance of nonlinear dynamic systems with uncertain external disturbances [11], [12], [14], [15]. However, the robust H_∞ control design of nonlinear dynamic system with external disturbance needs to solve a very complicated nonlinear partial differential Hamilton Jacobi Isaacs Equation (HJIE), which can not be efficiently solved analytically and numerically at present [11], [14], [15]. Therefore, several approximation methods have been proposed to interpolate some local linearized systems by fuzzy interpolation method [16], [17], gain scheduling method [21], [22] and global linearization [18]–[20] to approximate a nonlinear dynamic system so that the HJIE could be approximated by a set of local Riccati-like equations, which can be transformed to a set of linear matrix inequalities (LMIs) [20]. Then these interpolation methods for the robust H_∞ control design problem of nonlinear dynamic system with external disturbance can be transformed to the problem of how to solve a set of LMIs with the help of LMI toolbox in Matlab. The shortages of these interpolation methods are mentioned as follows: (i) In the process of transferring HJIE to a set of LMIs, we have performed the operation of inequality several times, leading to a very conservative result, (ii) a specific quadratic Lyapunov function $V(x(t)) = x^T(t)Px(t)$ for some $P = P^T > 0$ is selected as the solution of HJIE, which will limit the domain of the solution to HJIE and lead to a conservative result, and (iii) the state feedback control law will be the interpolation of N local control laws. More computations are needed for control law, especially for complex nonlinear system. Further, if the system state $x(t)$ is unavailable and to be estimated from output measurement, then the complex interpolatory observer-based output feedback controller is needed to compute at every time instant for the interpolatory observer-based output feedback control law. Obviously, the computation loadings of these H_∞ observer-based output feedback control laws are heavy and will prevent their practical applications, especially for highly nonlinear systems like quadrotor UAV systems [26].

Recently, a DNN-based H_∞ control scheme has been proposed for the stabilization control design problem of nonlinear time-varying dynamic systems with external disturbance [27]. The HJIE of robust H_∞ state feedback control design problem can be directly solved by the DNN-based learning algorithm so that the nonlinear H_∞ state feedback controller of nonlinear time-varying system can be obtained without the interpolation of local linear controllers by the conventional interpolation methods. However, in practical applications, the state variables of nonlinear dynamic system are not always all available and reference tracking control design is more appealing. These state variables can be only estimated from output measurement with the corruption of measurement noise. In this situation, the H_∞ observer-based output feedback control is more suitable to treat the robust reference tracking control of nonlinear dynamical systems with external disturbance and measurement noise in practical applications. Therefore, we need to solve a control-observer-coupled HJIE for robust H_∞ observer-based output feedback reference tracking control design problem of nonlinear stochastic systems under external disturbance and measurement noise. To avoid solving the very complicated control-observer-coupled HJIE for the H_∞ observer-based output feedback control design, T-S fuzzy observer-based state feedback control was proposed to interpolate a set of N^2 local linear observer-based controllers to achieve the robust H_∞ estimation and control performance by solving a set of N^2 control-observer-coupled LMIs by a so-called two-step design procedure [16]. These local interpolation methods need much effort to solve the robust H_∞ observer-based output feedback control problem. Further, it needs much computational time to calculate N^2 local observer-based output feedback control laws at every time instant for nonlinear dynamic system with external disturbance and output measurement noise. Since output feedback

tracking control is more useful to achieve a desired reference tracking design in practical applications, it is more appealing to design a robust H_∞ Luenberger observer-based reference tracking control design of nonlinear stochastic system with output measurement under external disturbance and measurement noise. Therefore, to avoid design complexity and save computation loading in the above interpolation methods, in this study, an HJIE-embedded DNN learning method is proposed as a co-design of H_∞ observer-based output feedback reference tracking control scheme and DNN learning algorithm for the robust H_∞ Luenberger observer-based reference tracking control design of nonlinear stochastic system with output measurement under external disturbance and measurement noise.

In this work, based on the augmented stochastic estimation error dynamic system and time-varying reference tracking error system, the minmax stochastic H_∞ Nash game strategy [15] is employed to minimize the worst-case effect of external disturbance and measurement noise on the state estimation error and reference tracking error. The robust H_∞ observer-based output feedback reference tracking control design problem of nonlinear stochastic systems needs to solve $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T}$ of control-observer-coupled HJIE, where $V(\tilde{x}(t), e(t), t)$ is the Lyapunov function of the estimation error $\tilde{x}(t)$ of observer, the tracking error $e(t)$ of controller, and time t for the design of control law $u^*(t)$ and observer gain $L^*(\hat{x}(t))$ of the robust H_∞ observer-based output feedback reference tracking control design. The difficulties in the design procedure of H_∞ observer-based output feedback reference tracking control of nonlinear stochastic system comprise: (i) The separation principle of observer and controller can not hold for nonlinear stochastic systems with external disturbance and output measurement noise so that HJIE of H_∞ observer-based output feedback reference tracking control design is control-observer-coupled; (ii) the HJIE is also a function of state $x(t)$, which is unavailable in our nonlinear stochastic system with output measurement. To overcome this difficulty, both observer dynamic system model and estimation error dynamic system model are needed to generate $\hat{x}(t)$ and $\tilde{x}(t)$, respectively, to provide $x(t) = \hat{x}(t) + \tilde{x}(t)$ for solving HJIE as shown in Fig. 1; (iii) in the off-line training phase, external disturbance $v(t)$ and measurement noise $n(t)$ are unavailable to generate output measurement $y(t)$ by nonlinear stochastic system model for observer to estimate system state. In this study, the worst-case external disturbance and measurement noise $v^*(t)$ and $n^*(t)$ of H_∞ observer-based reference tracking control strategy are employed to replace $v(t)$ and $n(t)$, respectively, and are inputed to nonlinear stochastic system model with H_∞ control law $u^*(t)$ to produce the measurement signal $y(t)$ for the off-line training as shown in Fig. 1; (iv) the most difficulty work is to solve $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T}$ from HJIE directly for H_∞ control law $u^*(t)$ and observer gain $L^*(\hat{x}(t))$. In this study, HJIE-emdedded DNN is employed to efficiently approach to $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T}$ via Adam learning algorithm [23]. We can prove that as the learning error of the embedded HJIE by Adam learning algorithm approaches to zero, the DNN with input $\tilde{x}(t)$ and $e(t)$ will output $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T}$, i.e., the HJIE-embedded DNN observer-based reference tracking control scheme in Fig. 1 can approach to the H_∞ observer-based output feedback reference tracking control design of nonlinear stochastic system with external disturbance and measurement noise. We have also proved that if the external disturbance and measurement noise are of finite energy, the mean-square asymptotical estimation to the true state (i.e., $\hat{x}(t) \to x(t)$) and the mean-square asymptotical tracking ability to the desired target $r(t)$ (i.e., $x(t) \to r(t)$) are both guaranteed by the proposed HJIE-embedded DNN observer-based reference tracking control scheme too.

In this study, the theoretical control-observer-coupled HJIE of H_∞ observer-based output feedback reference tracking control strategy could be employed as an expert knowledge of DNN-based H_∞ observer-based reference tracking control scheme. Based on the system models and embedded control-observer-coupled HJIE to train DNN to approach H_∞ observer-based reference tracking control strategy, we could not only apply DNN learning scheme to approach to the traditional complicated H_∞ nonlinear estimator-based output feedback reference tracking control design but also significantly reduce a larger amount of training data and training time than the conventional deep learning approaches via big data training.

The major contributions of this work are described as follows:

1) A novel DNN-based robust H_∞ observer-based reference tracking control scheme is proposed for nonlinear stochastic system with external disturbance and output measurement noise. The off-line training process of DNN can be accomplished by deep learning algorithm to approach to $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T}$ to solve the control-observer-coupled HJIE for the H_∞ control law, H_∞ observer gain and the worst-case external disturbance and measurement noise simultaneously to achieve the H_∞ observer-based output feedback tracking control design of nonlinear stochastic system with external disturbance and output measurement noise.
2) We could show that the proposed DNN-based output feedback tracking control scheme can approach to the theoretical H_∞ observer-based output feedback reference tracking control of nonlinear stochastic system by Adam learning algorithm. We could also prove that if the stochastic nonlinear system is with finite-energy external disturbance and measurement noise, the proposed DNN-based robust H_∞ observer-based output feedback tracking scheme will achieve the mean-square asymptotical state estimation, asymptotical reference tracking and asymptotical stability of closed loop nonlinear stochastic observer-based control system simultaneously.

3) By the proposed HJIE-embedded DNN H_∞ observer-based output feedback reference tracking control scheme, the traditional nonlinear stochastic system models and theoretical results of robust H_∞ observer-based output feedback tracking control could be employed to complement the traditional big data-driven deep learning approaches with more wide applications to nonlinear system control designs. Further, since system model, observer model and tracking error model are employed to help training DNN-based H_∞ observer-based output feedback reference tracking control scheme, we can save much training data and time than the conventional big data-driven deep learning approaches.

The remainder of this study is organized as follows: In section II, we will discuss the robust H_∞ observer-based output feedback tracking control design problem of nonlinear stochastic system with external disturbance and measurement noise. In section III, a novel control-observer-coupled HJIE-embedded H_∞ DNN observer-based reference tracking scheme is introduced to deal with output feedback reference tracking control design problem of a nonlinear stochastic system under external disturbance and measurement noise. In section IV, a numerical simulation of UAV reference tracking control design through output measurement is given to illustrate the design procedure and validate the reference tracking performance. Finally, a conclusion is given in section V.

Notation: A^T denotes the transpose of vector or matrix A; $P = P^T \geq 0$ denotes semi-positive definite matrix P; $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial [\tilde{x}^T(t)\ e^T(t)\ t]^T} \triangleq \begin{bmatrix} \frac{\partial V(\tilde{x}(t),e(t),t)}{\partial \tilde{x}(t)} \\ \frac{\partial V(\tilde{x}(t),e(t),t)}{\partial e(t)} \\ \frac{\partial V(\tilde{x}(t),e(t),t)}{\partial t} \end{bmatrix}$; $\mathbb{R}^n$ denotes the set of n-tuple real vectors; $\mathcal{L}_2[0,\infty)$ denotes a set of real functions $x(t) \in \mathbb{R}^n$ with finite energy, i.e., $x(t) \in \mathcal{L}_2[0,\infty)$, if $\|x(t)\|_2 = (E\int_0^\infty x^T(t)x(t)\,dt)^{\frac{1}{2}} < \infty$; I_a denotes the identity matrix with dimension $a \times a$; $diag\{A, B\} \triangleq \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$.

II. PROBLEM DESCRIPTION

For decades, H_∞ robust control designs have been the important research topics for a broad spectrum of application areas and impacts [10]–[15]. In practical applications, external disturbances and measurement noises are always unavoidable in real control systems; for example, the loadings and environment interferences are considered as external disturbance, and output measurement noises always occur when state variables are unavailable and we need to estimate state variables from output measurement for state feedback control. During past decades, robust H_∞ observer-based output feedback control strategies have been developed for efficient attenuation of uncertain external disturbance and measurement noise on nonlinear quadratic stabilization and nonlinear state estimation of nonlinear dynamic system from the minmax Nash game point of view [16], [17]. In this section, we review the robust H_∞ observer-based reference tracking control design of nonlinear stochastic system with external disturbance and output measurement noise.

Consider the following nonlinear stochastic system with external disturbance and output measurement

$$\begin{aligned} \dot{x}(t) &= F(x(t)) + G(x(t))u(t) + D(x(t))v(t),\ x(0) = x_0 \\ y(t) &= C(x(t)) + n(t) \end{aligned} \tag{1}$$

where $x(t) \in \mathbb{R}^n$ is the state vector, $x_0 \in \mathbb{R}^n$ denotes the initial condition, $u(t) \in \mathbb{R}^m$ is the control input, $y(t) \in \mathbb{R}^l$ is the measurement output, $F(x(t)) \in \mathbb{R}^n$, $G(x(t)) \in \mathbb{R}^{n\times m}$, $C(x(t)) \in \mathbb{R}^l$ and $D(x(t)) \in \mathbb{R}^{n\times k}$ are system functions. These system functions are assumed to satisfy with the Lipschitz continuity. $v(t) \in \mathbb{R}^k$ and $n(t) \in \mathbb{R}^l$ denote the random external disturbance and measurement noise, respectively. In the nonlinear stochastic system in (1), only output measurement $y(t)$ is available.

In this study, we want to control the state $x(t)$ of nonlinear stochastic system in (1) from output $y(t)$ to track a desired reference target $r(t)$. Let us denote the reference tracking error as $e(t) = x(t) - r(t)$. Then the reference tracking error dynamic can be described as follows:

$$\begin{aligned} \dot{e}(t) &= \dot{x}(t) - \dot{r}(t) \\ &= F(x(t)) + G(x(t))u(t) + D(x(t))v(t) - \dot{r}(t) \\ &= F(e(t) + r(t)) - \dot{r}(t) + G(e(t) + r(t))u(t) \\ &\quad + D(e(t) + r(t))v(t) \end{aligned} \tag{2}$$

Let us denote

$$\begin{aligned} F_e(e(t), t) &\triangleq F(e(t) + r(t)) - \dot{r}(t) \\ G_e(e(t), t) &\triangleq G(e(t) + r(t)) \\ D_e(e(t), t) &\triangleq D(e(t) + r(t)) \end{aligned} \tag{3}$$

Then we get the reference tracking error time-varying dynamic equation as

$$\dot{e}(t) = F_e(e(t), t) + G_e(e(t), t)u(t) + D_e(e(t), t)v(t) \tag{4}$$

The design purpose of this study is to specify a robust state feedback reference tracking control for the nonlinear stochastic system in (1). However, the state $x(t)$ in (1) is unavailable. Therefore, the following Luenberger observer is employed to estimate the state vector for observer-based output feedback reference tracking control of nonlinear stochastic system in (1)

$$\begin{aligned} \dot{\hat{x}}(t) &= F(\hat{x}(t)) + G(\hat{x}(t))u(t) + L(\hat{x}(t))(y(t) - C(\hat{x}(t))) \\ u(t) &= K(\hat{x}(t), e(t)) \end{aligned} \tag{5}$$

where $L(\hat{x}(t)) \in \mathbb{R}^{n\times l}$ denotes the observer gain and $K(\hat{x}(t), e(t))$ is the control gain based on the state estimate $\hat{x}(t)$ and the tracking error $e(t)$.

Based on the above analysis, the following H_∞ observer-based output feedback reference tracking control design problem for the nonlinear stochastic system in (1) is how to specify control gain $K(\hat{x}(t), e(t))$ and observer gain $L(\hat{x}(t))$

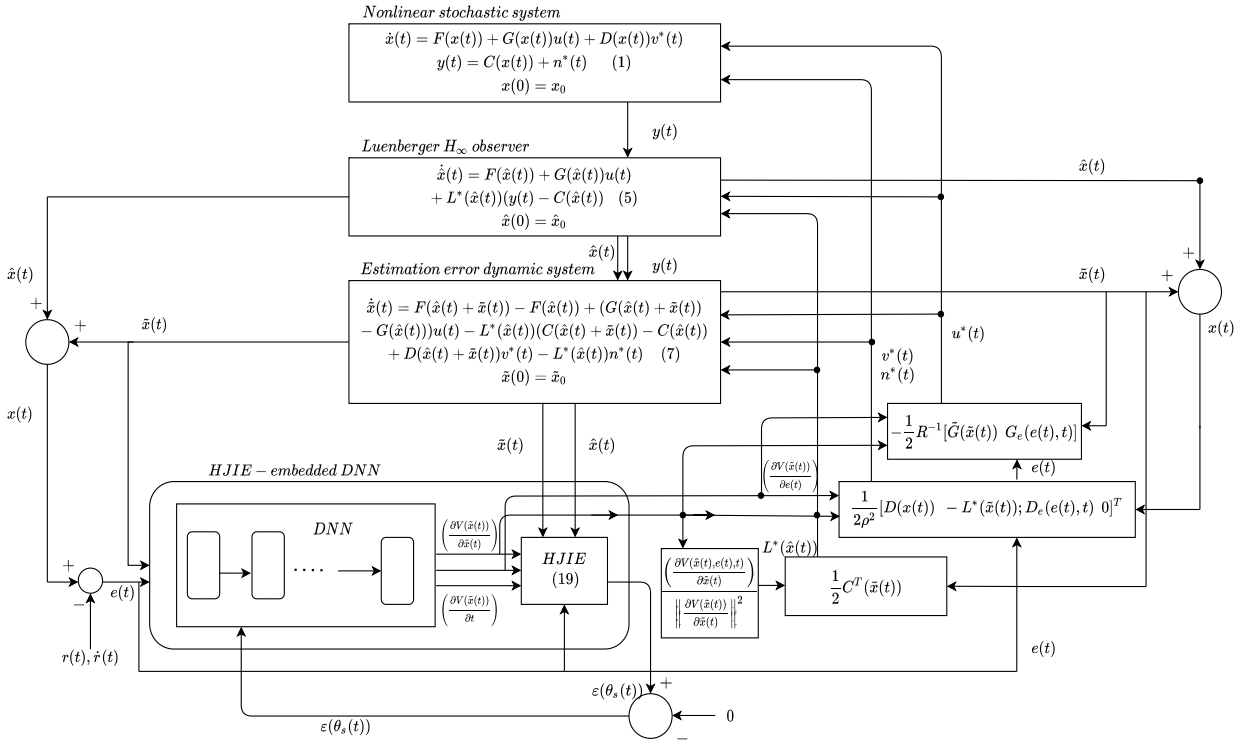

FIGURE 1. The flow chart of the HJIE-embedded DNN-based H_∞ observer-based output feedback reference tracking control scheme of nonlinear stochastic system in (1) with external disturbance and measurement noise. Since the real random external disturbance $v(t)$ and measurement noise $n(t)$ are unavailable in the off-line training phase, the worst-case external disturbance $v^*(t)$ and measurement noise $n^*(t)$, which are generated by the output $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)e^T(t)]^T}$ of DNN based on (13), are used to replace the real $v(t)$ and $n(t)$ to generate $y(t)$ by the the stochastic system model with control input $u^*(t)$, which is also generated by the output $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)]^T}$ of DNN based on (14). Then $y(t)$ is inputed to Luenberger observer in (5) to generate $\hat{x}(t)$ with observer gain $L^*(\hat{x}(t))$ being generated by the output $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)]^T}$ of DNN based on (15). Further $u^*(t)$, $L^*(\hat{x}(t))$ and $\begin{bmatrix} v^*(t) \\ n^*(t) \end{bmatrix}$, which are also generated by the $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)]^T}$ of DNN, are also inputed to the estimation error system in (7) to generate estimation error $\tilde{x}(t)$. Then we can obtain the state $x(t) = \hat{x}(t) + \tilde{x}(t)$ to obtain the tracking error by $e(t) = x(t) - r(t)$. Finally, the state estimation error $\tilde{x}(t)$ and reference tracking error $e(t)$ are both inputed to DNN to be expected to output $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T}$ after the proposed HJIE-embedded deep learning training by Adam learning algorithm in the off-line training phase. In the on-line operation phase, the flow chart of HJIE-embedded DNN-based H_∞ observer-based output feedback reference control scheme is similar to the off-line training phase except the Adam learning algorithm being stopped and $v^*(t)$ and $n^*(t)$ being replaced by the real $v(t)$ and $n(t)$, respectively.

in (5) such that the worst-case effect of external disturbance $v(t)$ and measurement noise $n(t)$ on the estimation error and reference tracking error must be minimized and below a prescribed attenuation level ρ as follows [11], [15]:

$$\min_{\substack{K(\hat{x}(t),e(t)) \\ L(\hat{x}(t))}} \max_{\substack{v(t)\in\mathcal{L}_2[0,t_f] \\ n(t)\in\mathcal{L}_2[0,t_f]}} \frac{\begin{array}{c} E\{\int_0^{t_f}((x(t)-\hat{x}(t))^T Q_1 \\ \times(x(t)-\hat{x}(t)) + e^T(t)Q_2 e(t) \\ +u^T(t)Ru(t))dt\} \end{array}}{E\{\int_0^{t_f}(v^T(t)v(t) + n^T(t)n(t))dt\}} \le \rho^2 \tag{6}$$

where $E\{\cdot\}$ denotes the expectation of $\{\cdot\}$, t_f denotes the terminal time, the weighting matrices $Q_1 \in \mathbb{R}^{n\times n}, Q_2 \in \mathbb{R}^{n\times n}$ and $R \in \mathbb{R}^{m\times m}$ are specified by control designer as the tradeoff among the estimation error, tracking error and control effort. The physical meaning of (6) is that the worst-case effect of external disturbance and measurement noise on the state estimation error, reference tracking error and control effort should be minimized by observer-based control $u(t) = K(\hat{x}(t), e(t))$ and observer gain $L(\hat{x}(t))$ and must be less than a prescribed level ρ simultaneously.

From (1) and (5), we could obtain the state estimation error equation of $\tilde{x}(t) = x(t) - \hat{x}(t)$ as follows:

$$\begin{aligned} \dot{\tilde{x}}(t) &= F(x(t)) - F(\hat{x}(t)) + (G(x(t)) - G(\hat{x}(t)))u(t) \\ &\quad -L(\hat{x}(t))(C(x(t)) - C(\hat{x}(t))) \\ &\quad +D(x(t))v(t) - L(\hat{x}(t))n(t) \\ &= \tilde{F}(\tilde{x}(t)) + \tilde{G}(\tilde{x}(t))u(t) - L(\hat{x}(t))\tilde{C}(\tilde{x}(t)) \\ &\quad + \begin{bmatrix} D(x(t)) & -L(\hat{x}(t)) \end{bmatrix} \bar{v}(t) \end{aligned} \tag{7}$$

where $\tilde{F}(\tilde{x}(t)) = F(x(t)) - F(\hat{x}(t))$, $\tilde{G}(\tilde{x}(t)) = G(x(t)) - G(\hat{x}(t))$ and $\tilde{C}(\tilde{x}(t)) = C(x(t)) - C(\hat{x}(t))$.

Then the H_∞ observer-based output feedback reference tracking control strategy in (6) could be reformulated

as follows:

$$\min_{\substack{K(\hat{x}(t),e(t))\\ L(\hat{x}(t))}} \max_{\substack{v(t)\in\mathcal{L}_2[0,t_f]\\ n(t)\in\mathcal{L}_2[0,t_f]}} \frac{E\{\int_0^{t_f}(\tilde{x}(t)^T Q_1 \tilde{x}(t) + e^T(t)Q_2 e(t) + u^T(t)Ru(t))dt\}}{E\{\int_0^{t_f}(v^T(t)v(t)+n^T(t)n(t))dt\}} \le \rho^2$$

or

$$\min_{\substack{K(\hat{x}(t),e(t))\\ L(\hat{x}(t))}} \max_{\bar{v}(t)\in\mathcal{L}_2[0,t_f]} \frac{E\{\int_0^{t_f}(\bar{x}^T(t)\bar{Q}\bar{x}(t) + u^T(t)Ru(t))dt\}}{E\{\int_0^{t_f}\bar{v}^T(t)\bar{v}(t)dt\}} \le \rho^2 \quad (8)$$

where

$$\bar{x}(t) = [\,\tilde{x}^T(t)\ e^T(t)\,]^T$$
$$\bar{Q} = diag\{Q_1, Q_2\}$$
$$\bar{v}(t) = [\,v^T(t)\ n^T(t)\,]^T$$

In (8), to simplify the design procedure, the robust H_∞ observer-based output feedback reference tracking control strategy in (8) becomes the robust H_∞ stabilization strategy of the following augmented nonlinear stochastic time-varying error system

$$\dot{\bar{x}}(t) = \begin{bmatrix} \tilde{F}(\tilde{x}(t)) - L(\hat{x}(t))\tilde{C}(\tilde{x}(t)) \\ F_e(e(t),t) \end{bmatrix} + \begin{bmatrix} \tilde{G}(\tilde{x}(t)) \\ G_e(e(t),t) \end{bmatrix} \times u(t) + \begin{bmatrix} D(x(t)) & -L(\hat{x}(t)) \\ D_e(e(t),t) & 0 \end{bmatrix} \bar{v}(t) \quad (9)$$

In the robust H_∞ stabilization strategy in (8), we assume the initial conditions $\tilde{x}(0)$ and $e(0)$ are all 0. If $\tilde{x}(0) \neq 0$ and $e(0) \neq 0$, the effect of initial energy $V(\tilde{x}(0), e(0), 0)$ due to the initial condition should be extracted as follows:

$$\min_{\substack{K(\hat{x}(t),e(t))\\ L(\hat{x}(t))}} \max_{\bar{v}(t)\in\mathcal{L}_2[0,t_f]} \frac{E\{\int_0^{t_f}(\bar{x}^T(t)\bar{Q}\bar{x}(t) + u^T(t)Ru(t))dt - V(\tilde{x}(0), e(0), 0)\}}{E\{\int_0^{t_f}\bar{v}^T(t)\bar{v}(t)dt\}} \le \rho^2 \quad (10)$$

where $V(\tilde{x}(t), e(t), t)$ denotes the Lyapunov energy function of the nonlinear augmented stochastic time-varying error system in (9).

Based on the above analysis, the complex H_∞ observer-based reference tracking control design problem in (6), which involves the nonlinear stochastic system in (1), tracking error time-varying dynamic in (4) and observer-based output feedback in (5), becomes how to solve the minmax H_∞ stabilization design problem in (10) for the nonlinear augmented stochastic time-varying error system in (9). This control design problem reformulation will significantly simplify the control design procedure in the sequel.

In (10), one player $\bar{v}(t)$ wants to maximize the payoff function while other players $K(\hat{x}(t), e(t))$ and $L(\hat{x}(t))$ try to minimize the payoff function. However, it is not easy to directly solve the minmax game problem of the fractional payoff function in (10) for the nonlinear stochastic augmented error system in (9). The following indirect two-step method [11] is employed to solve minmax Nash game problem in (10). Since the selection of the player $\bar{v}(t)$ in the denominator of payoff function in (10) is independent on $K(\hat{x}(t), e(t))$ and $L(\hat{x}(t))$, the minmax game problem in (10) is equivalent to [11], [15]

$$\min_{\substack{K(\hat{x}(t),e(t))\\ L(\hat{x}(t))}} \max_{\bar{v}(t)\in\mathcal{L}_2[0,t_f]} E\{\int_0^{t_f}(\bar{x}^T(t))\bar{Q}\bar{x}(t)) + u^T(t)Ru(t) - \rho^2\bar{v}^T(t))\bar{v}(t))dt\} \le E\{V(\tilde{x}(0), e(0), 0)\} \quad (11)$$

Then a two-step method is employed to solve the constrained minmax Nash quadratic game problem in (11), i.e., (i) in the first step, we need to solve the following minmax Nash quadratic game problem at first

$$J = \min_{\substack{K(\hat{x}(t),e(t))\\ L(\hat{x}(t))}} \max_{\bar{v}(t)\in\mathcal{L}_2[0,t_f]} E\{\int_0^{t_f}(\bar{x}^T(t)\bar{Q}\bar{x}(t) + u^T(t)Ru(t) - \rho^2\bar{v}^T(t)\bar{v}(t))dt\} \quad (12)$$

then (ii) in the second step, we need to guarantee $J \le E\{V(\tilde{x}(0), e(0), 0)\}$.

Based on the above two-step indirect method, the robust H_∞ observer-based output feedback reference tracking control strategy in (10) for the augmented time-varying stochastic error system in (9) is solved by the following theorem.

Theorem 1: (a) For the nonlinear stochastic system in (1) with external disturbance and output measurement noise, suppose the Luenberger observer-based control law in (5) is employed to achieve the robust H_∞ observer-based output feedback reference tracking control strategy in (6), (10) or (11). Then the worst-case external disturbance $v^*(t)$ and measurement noise $n^*(t)$ as well as H_∞ control gain $K^*(\hat{x}(t), e(t))$ and observer gain $L^*(\hat{x}(t))$ are given as follows:

$$\bar{v}^*(t) = \begin{bmatrix} v^*(t) \\ n^*(t) \end{bmatrix} = \frac{1}{2\rho^2}\begin{bmatrix} D(x(t)) & -L^*(\hat{x}(t)) \\ D_e(e(t),t) & 0 \end{bmatrix}^T \times \left(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)]^T}\right) \quad (13)$$

$$u^*(t) = K^*(\hat{x}(t), e(t)) = -\frac{1}{2}R^{-1}\begin{bmatrix} \tilde{G}(\tilde{x}(t)) \\ G_e(e(t),t) \end{bmatrix}^T \left(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)]^T}\right) \quad (14)$$

$$L^*(\hat{x}(t)) = \frac{1}{2}\frac{\left(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial\tilde{x}(t)}\right)}{\left\|\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial\tilde{x}(t)}\right\|^2}\tilde{C}^T(\tilde{x}(t)) \quad (15)$$

where the Lyapunov function $V(\tilde{x}(t), e(t), t) > 0$ with $V(0, 0, t) = 0$ is the solution of the following time-varying HJIE

$$HJIE = \frac{\partial V(\tilde{x}(t), e(t), t)}{\partial t} + \bar{x}^T(t)\bar{Q}\bar{x}(t) - \frac{1}{4}\left(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)]^T}\right)^T \bar{G}(\tilde{x}(t), e(t), t)R^{-1} \times \bar{G}^T(\tilde{x}(t), e(t), t)\left(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)]^T}\right)$$

$$+(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)]^T})^T\bar{F}(\tilde{x}(t),e(t),t)$$
$$+\frac{1}{4\rho^2}(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)]^T})^T\bar{D}(\tilde{x}(t),e(t),t)$$
$$\times(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)]^T})-\frac{1}{16\rho^2}\tilde{C}^T(\tilde{x}(t))\tilde{C}(\tilde{x}(t))=0 \tag{16}$$

where

$$\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)]^T}\triangleq\begin{bmatrix}\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial\tilde{x}(t)}\\ \frac{\partial V(\tilde{x}(t),e(t),t)}{\partial e(t)}\end{bmatrix}$$
$$\bar{G}(\tilde{x}(t),e(t),t)=\begin{bmatrix}\tilde{G}(\tilde{x}(t))\\ G_e(e(t),t)\end{bmatrix}$$
$$\bar{F}(\tilde{x}(t),e(t),t)=\begin{bmatrix}\tilde{F}(\tilde{x}(t))\\ F_e(e(t),t)\end{bmatrix}$$
$$\bar{D}(\tilde{x}(t),e(t),t)=\begin{bmatrix}D(x(t))D^T(x(t)) & D(x(t))D_e^T(e(t),t)\\ D_e(e(t),t)D^T(x(t)) & D_e(e(t),t)D_e^T(e(t),t)\end{bmatrix}$$

(b) In the nonlinear stochastic system in (1), if external disturbance $v(t)\in\mathcal{L}_2[0,\infty)$ and measurement noise $n(t)\in\mathcal{L}_2[0,\infty)$, then the asymptotical estimation and tracking ability will be achieved by the proposed robust H_∞ observer-based output feedback control in (5) with control gain $K^*(\hat{x}(t),e(t))$ and observer gain $L^*(\hat{x}(t))$ in (a), i.e., $\tilde{x}(t)\to 0$ and $e(t)\to 0$ as $t_f\to\infty$. Further, if $r(t)\in\mathcal{L}_2[0,\infty)$, then the mean-square asymptotical stability of the closed loop system is also guaranteed.

Proof: See the Appendix A. □

From (16), all terms of HJIE contain the estimation error $\tilde{x}(t)$, system state $x(t)$ and tracking error $e(t)=x(t)-r(t)$, i.e., HJIE in (16) is a control-observer-coupled HJIE. It is very difficult to solve HJIE in (16) analytically or numerically, especially, when $x(t)$ is unavailable.

Remark 1: In general, for the H_∞ observer-based output feedback tracking control design, there are two-coupled HJIEs to be solved, one for observer and another for controller. However, when the state estimation error and reference tracking error are simultaneously considered in the H_∞ observer-based output feedback reference tracking performance in (6) or (8) and the state estimation error equation and reference tracking error equation are combined as the augmented nonlinear error system in (9), we could obtain a control-observer-coupled HJIE in (16) for H_∞ observer-based output feedback tracking control strategy. This will simplify the design procedure of DNN-based H_∞ observer-based reference tracking scheme.

In general, it is very difficult to solve the partial time-varying nonlinear differential equation HJIE in (16) for the H_∞ observer-based reference tracking control design of $K^*(\hat{x}(t),e(t))$ in (14) and $L^*(\hat{x}(t))$ in (15) for nonlinear stochastic system in (1) with external disturbance and measurement noise. In the last decades, several approximation methods like the global linearization method [18], [19], T-S fuzzy interpolation method [16] and gain scheduling method [21], [22], etc., have been used to interpolate N local stochastic linearized systems to approximate the nonlinear system with output measurement in (1). Then a set of N^2 local linear observer-based state feedback controllers are interpolated to approximate the nonlinear observer-based output feedback controller in (5). Based on these interpolation approximation methods and with the choice of quadratic Lyapunov function $V(\tilde{x}(t),e(t),t)=\bar{x}^T(t)P\bar{x}(t)$, the HJIE in (16) will become a set of N^2 control gains $\{K_i\}_{i=1}^N$-observer gains $\{L_j\}_{j=1}^N$-coupled Riccati-like equations, which could be transformed to N^2 control gains-observer gains-coupled bilinear matrix inequalities (BMIs) [16]. Finally, a two-step method is used to solve these complex control gains-observer gains-coupled BMIs. The shortages of these interpolation methods are given as follows: In some practical nonlinear systems, in order to reduce the approximation error, there are a large number of local control gains $\{K_i\}_{i=1}^N$ and observer gains $\{L_j\}_{j=1}^N$ to be solved from N^2 coupled BMIs. For example, based on T-S fuzzy approximation method, there are $\{K_i\}_{i=1}^{125}$ and $\{L_j\}_{j=1}^{125}$ to be solved from 15625 K_i and L_j-coupled BMIs in the quadrotor UAV H_∞ observer-based tracking control design [17]. It is a very complex design procedure to solve these $\{K_i\}_{i=1}^{125}$ and $\{L_j\}_{j=1}^{125}$ from 15625 K_i and L_j-coupled BMIs by the so-called two-step design procedure for robust H_∞ fuzzy observer-based quadrotor UAV tracking design problem. Further, we need to compute the following T-S fuzzy observer-based state feedback control $u(t)$ at every time instant [17]

$$\dot{\hat{x}}(t)=\sum_{i=1}^{125}\sum_{j=1}^{125}h_i(\hat{x}(t))h_j(\hat{x}(t))(A_i\hat{x}(t)-B_iu(t))$$
$$+L_i(y(t)-C_j(\hat{x}(t)))$$
$$u(t)=\sum_{i=1}^{125}h_i(\hat{x}(t))K_i\hat{x}(t) \tag{17}$$

where $h_i(\hat{x}(t)), i=1,\ldots,125$ are the complicated fuzzy interpolation bases. Moreover, the selection of quadratic Lyapunov solution $V(\tilde{x}(t),e(t),t)=\bar{x}^T(t)P\bar{x}(t)$ for HJIE in (16) is very conservative, which will limit the solution of nonlinear $V(\tilde{x}(t),e(t),t)$ in (16).

In the above minmax H_∞ observer-based output feedback reference tracking control design problem of nonlinear stochastic system in (1), the main difficulty lies in how to solve the two partial differential $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial t}$ and $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)]^T}$ from HJIE in (16) for $u^*(t)$ in (14) and $L^*(\hat{x}(t))$ in (15). In this study, an HJIE-embedded deep learning approach in Fig. 1 and Fig. 2 will be proposed

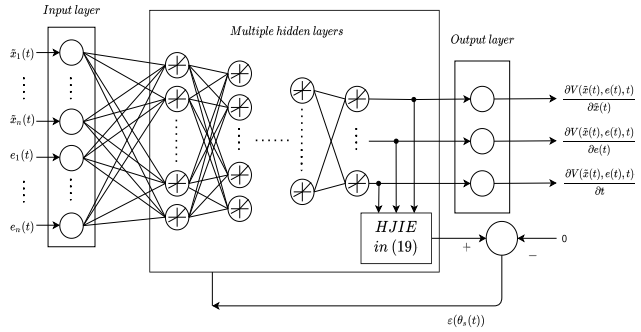

FIGURE 2. The architecture of the HJIE-embedded DNN is to be trained by Adam learning algorithm to solve $HJIE = 0$ in (19) in the off-line training phase. After Adam learning algorithm approaches to 0, the HJIE-embedded DNN with input $\begin{bmatrix} \tilde{x}(t) \\ e(t) \end{bmatrix}$ can output $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)]^T}$ to generate the robust H_∞ control $u^*(t)$ in (14), H_∞ observer gain $L^*(\hat{x}(t))$ in (15) and worst-case $\begin{bmatrix} v^*(t) \\ n^*(t) \end{bmatrix}$ in (13) simultaneously. Then Luenberger observer in (5) can generate $\hat{x}(t)$ and the estimation error equation in (7) can generate $\tilde{x}(t)$, so that $x(t) = \hat{x}(t) + \tilde{x}(t)$ can be obtained for tracking error $e(t) = x(t) - r(t)$. Finally, $\tilde{x}(t)$ and $e(t)$ are fedback as input to DNN to begin another cycle of Adam learning process based on the error $\varepsilon(\theta(t))$ of $HJIE_\varepsilon$ in (20).

to directly solve control-observer-coupled HJIE in (16) for the H_∞ observer-based output feedback reference tracking control design problem of nonlinear stochastic system with external disturbance and output measurement noise in (1).

III. HJIE-EMBEDDED DNN-BASED H_∞ OBSERVER-BASED OUTPUT FEEDBACK REFERENCE TRACKING CONTROL DESIGN OF NONLINEAR STOCHASTIC SYSTEMS

For the nonlinear stochastic system with random external disturbance and output measurement noise in (1), in this study, the minmax H_∞ observer-based output feedback reference tracking control strategy in (6) is employed by the observer-based output feedback control in (5) to minimize the worst-case effect of external disturbance $v^*(t)$ and measurement noise $n^*(t)$ on the state estimation error $\tilde{x}(t)$ and reference tracking error $e(t)$ as well as control effort $u(t)$ from the energy perspective. From Theorem 1, we need to solve the very complicated time-varying partial differential HJIE in (16) for $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial t}$ and $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)]^T}$ to obtain $K^*(\hat{x}(t), e(t))$ in (14) and $L^*(\hat{x}(t))$ in (15) for nonlinear observer-based output feedback control law in (5) and $\bar{v}^*(t)$ in (13) for the worst-case external disturbance and measurement noise. In this study, in order to avoid the above complex approximation methods of interpolation through local linearization schemes, we employ an HJIE-embedded DNN to approach $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial t}$ and $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)]^T}$ directly instead of $V(\tilde{x}(t), e(t), t)$ in the conventional methods [14]. The reason is that even $V(\tilde{x}(t), e(t), t)$ is solved [14], it is still very difficult to calculate $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)]^T}$ for $u^*(t)$ in (14), $L^*(\hat{x}(t))$ in (15) and $\bar{v}^*(t)$ in (13) in the real-time H_∞ observer-based output feedback tracking process.

For the convenience of design, let us denote

$$\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T} \triangleq \begin{bmatrix} \frac{\partial V(\tilde{x}(t),e(t),t)}{\partial \tilde{x}(t)} \\ \frac{\partial V(\tilde{x}(t),e(t),t)}{\partial e(t)} \\ \frac{\partial V(\tilde{x}(t),e(t),t)}{\partial t} \end{bmatrix} \tag{18}$$

Then HJIE in (16) can be rewritten as

$$\begin{aligned} HJIE =\ & \bar{x}^T(t)\bar{Q}\bar{x}(t) - \frac{1}{4}\left(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T}\right)^T \\ & \times \begin{bmatrix} \bar{G}(\tilde{x}(t), e(t), t) \\ 0 \end{bmatrix} R^{-1} \begin{bmatrix} \bar{G}(\tilde{x}(t), e(t), t) \\ 0 \end{bmatrix}^T \\ & \times \left(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T}\right) + \left(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T}\right)^T \\ & \times \begin{bmatrix} \bar{F}(\tilde{x}(t), e(t), t) \\ 1 \end{bmatrix} + \frac{1}{4\rho^2}\left(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T}\right)^T \\ & \times \begin{bmatrix} \bar{D}(\tilde{x}(t), e(t), t) & 0 \\ 0 & 0 \end{bmatrix} \left(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T}\right) \\ & - \frac{1}{16\rho^2}\tilde{C}^T(\tilde{x}(t))\tilde{C}(\tilde{x}(t)) = 0 \end{aligned} \tag{19}$$

Recently neural network is considered as an universal approximator to any complex function after an efficient deep learning approach [23], [24], [28]. Therefore, in this study, DNN is employed with deep learning approach to solve $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T}$ from HJIE in (19) for the H_∞ control law $u^*(t) = K^*(\hat{x}(t), e(t))$ in (14), the worst-case external disturbance and measurement noise $\begin{bmatrix} v^*(t) \\ n^*(t) \end{bmatrix}$ in (13) and H_∞ observer gain $L^*(\hat{x}(t))$ in (15) as shown in Fig. 1.

In the off-line training phase, since the external disturbance $v(t)$ and measurement noise $n(t)$ are unavailable for the nonlinear stochastic system in (1), the worst-case external disturbance $v^*(t)$ and measurement noise $n^*(t)$ generated by (13) are employed to replace $v(t)$ and $n(t)$ to generate $y(t)$ by nonlinear system model with the H_∞ control law $u^*(t)$ in (14), which does not affect the performance of H_∞ observer-based reference tracking control strategy in (6) because it is designed based on the worst-case $v^*(t)$ and $n^*(t)$ in (6). Since state variable $x(t)$ of HJIE in (19) is unavailable too, we need nonlinear Luenberger observer in (5) to generate $\hat{x}(t)$ and state estimation error equation in (7) to generate estimation error $\tilde{x}(t)$ so that system state $x(t)$ can be obtained by $x(t) = \hat{x}(t) + \tilde{x}(t)$ for $F(x(t))$, $G(x(t))$, $C(x(t))$ and $D(x(t))$ in (19), and therefore the tracking error $e(t)$ can be obtained by $e(t) = x(t) - r(t)$ for $F_e(e(t), t)$, $G_e(e(t), t)$ and $D_e(e(t), t)$ in (19) too. Then $\tilde{x}(t)$ and $e(t)$ are inputed to DNN to expect generating $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T}$ after deep learning scheme via the HJIE-embedded Adam learning algorithm. In this study, the output $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T}$ of DNN will be sent to the embedded HJIE to check the value of HJIE in (19). If the DNN output approaches to the expected $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T}$ of HJIE after the training of deep learning scheme, then HJIE in (19) will approach to 0 too. If not, DNN will output $\left(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T}\right)_\varepsilon$

which leads to the following result of HJIE

$$HJIE_\varepsilon = \bar{x}^T(t)\bar{Q}\bar{x}(t) - \frac{1}{4}(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T})_\varepsilon^T \times \begin{bmatrix} \bar{G}(\tilde{x}(t),e(t),t) \\ 0 \end{bmatrix} R^{-1} \begin{bmatrix} \bar{G}(\tilde{x}(t),e(t),t) \\ 0 \end{bmatrix}^T \times (\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T})_\varepsilon + (\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T})_\varepsilon^T \times \begin{bmatrix} \bar{F}(\tilde{x}(t),e(t),t) \\ 1 \end{bmatrix} + \frac{1}{4\rho^2}(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T})_\varepsilon^T \times \begin{bmatrix} \bar{D}(\tilde{x}(t),e(t),t) & 0 \\ 0 & 0 \end{bmatrix} (\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T})_\varepsilon - \frac{1}{16\rho^2}\tilde{C}^T(\tilde{x}(t))\tilde{C}(\tilde{x}(t)) = \varepsilon(\theta(t)) \quad (20)$$

Then $\varepsilon(\theta(t))$ will be fedback to train the weighting parameters of neurons in DNN as shown in Fig. 2 until the output of DNN approaches to $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T}$ as shown in the flow chart of Fig. 1. At the same time, the output $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)]^T}$ of DNN is multiplied by $-\frac{1}{2}R^{-1}\bar{G}^T(\tilde{x}(t),e(t),t)$ to generate the H_∞ control $u^*(t)$ in (14) and by $\frac{1}{2\rho^2}\begin{bmatrix} D(x(t)) & -L^*(\hat{x}(t)) \\ D_e(e(t),t) & 0 \end{bmatrix}^T$ to generate $\bar{v}^*(t)$ in (13). Further, the output $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial\tilde{x}(t)}$ is also used to produce $\frac{1}{2}\frac{\left(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial\tilde{x}(t)}\right)}{\left\|\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial\tilde{x}(t)}\right\|^2}\tilde{C}^T(\tilde{x}(t))$ for H_∞ observer gain $L^*(\hat{x}(t))$ in (15). These information of H_∞ control $u^*(t)$, H_∞ observer gain $L^*(\hat{x}(t))$ and the worst-case $\bar{v}^*(t)$ are sent to nonlinear system model to generate $y(t)$, Luenberger observer in (5) to generate $\hat{x}(t)$ and the estimation error equation in (7) to generate $\tilde{x}(t)$ and then generate $x(t) = \hat{x}(t) + \tilde{x}(t)$ to obtain $e(t) = x(t) - r(t)$. Finally, $\tilde{x}(t)$ and $e(t)$ are inputed to DNN for training and $\tilde{x}(t)$, $e(t)$ and $x(t)$ are also sent to HJIE to calculate its value $\varepsilon(\theta(t))$ for deep learning training as shown in the flow chart of Fig. 1 in the off-line training phase of HJIE-embedded DNN-based H_∞ observer-based output feedback tracking control scheme of nonlinear stochastic system with external disturbance and measurement noise in (1).

After the off-line training phase with $\varepsilon(\theta(t)) = 0$, DNN can output $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T}$. Then the proposed HJIE-embedded DNN-based H_∞ observer-based reference tracking control scheme is shifted to the on-line operation phase as shown in Fig. 1. In the on-line operation phase, the output $y(t)$ of nonlinear stochastic system is generated by the robust H_∞ observer-based state feedback control $u^*(t)$ in (14) and real $v(t)$ and $n(t)$. Therefore, we do not need $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)]^T}$ of DNN output to generate $\bar{v}^*(t)$ by (13) in the on-line operation phase. The remainder procedure is similar to the off-line training phase. In the on-line operation phase, generally, we do not need to train DNN again. However, in some situation, if $\varepsilon(\theta(t)) > \delta$ for a specified small value δ, we could also start on Adam learning algorithm without affecting the operation procedure of DNN-based H_∞ observer-based reference tracking control scheme in the on-line operation phase.

The architecture of HJIE-embedded DNN in the proposed robust H_∞ observer-based output feedback reference tracking control scheme consists of input layer, multiple hidden layers, HJIE layer and output layer as shown in Fig. 2. The estimation error $\tilde{x}(t)$ and reference tracking error $e(t)$ are inputed to DNN. Input layer consists of $2n$ nodes, n nodes for $\tilde{x}(t)$ and another n nodes for $e(t)$. There are $2n+1$ nodes in output layer, n nodes for outputing $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial\tilde{x}(t)}$, another n nodes for $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial e(t)}$ and one node for $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial t}$.

The neurons of DNN are with LeakyReLU $\sigma(x)$ as the activation function of x in hidden layers. LeakyReLU is not equal to zero but a constant gradient as input is negative, and is the same as ReLU while input is positive [23]. This way can keep the advantage of ReLU and avoid the problem of dead ReLU when input is negative, i.e., the operation of LeakyReLU is given as follows:

$$\sigma(x(t)) = \begin{cases} a_1 x(t), & \text{if } x(t) > 0 \\ a_2 x(t), & \text{if } x(t) \le 0 \end{cases}$$

where a_1 and a_2 are some constant with $a_1, a_2 \in (0, 1)$.

The error $\varepsilon(\theta(t))$ of HJIE in (20) will be fedback to DNN to train the weighting parameters and biases of neurons in DNN via the following Adam learning algorithm by minimizing the objective function $\varepsilon^2(\theta_s(t))$ [23], [24],

$$\theta_s(t) = \theta_{s-1}(t) - \frac{l}{\sqrt{\tilde{v}_s(t)+\zeta}}\tilde{m}_s(t), \quad s = 1, \cdots, S \quad (21)$$

where $\theta_s(t)$ is the weighting parameter vector of the sth training iteration at time t, which is to be trained to let DNN output $(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T})_\varepsilon$ at time t. l is the learning rate or stepsize and S is the number of training time steps at each time t. ζ is a small number to prevent the denominator from being 0. $\tilde{m}_s(t)$ and $\tilde{v}_s(t)$ are bias-corrected estimators defined as follows [23]:

$$\tilde{m}_s(t) = \frac{m_s(t)}{1-\mu_1^s}, \ \tilde{v}_s(t) = \frac{v_s(t)}{1-\mu_2^s} \quad (22)$$

where

$$m_s(t) = \mu_1 m_{s-1}(t) + (1-\mu_1)g_s(t)$$
$$v_s(t) = \mu_2 v_{s-1}(t) + (1-\mu_2)g_s^2(t) \quad (23)$$

where $g_s(t) = \frac{\partial}{\partial\theta_s(t)}\sqrt{\frac{1}{B}\sum^B \varepsilon^2(\theta_s(t))}$ is the gradient vector of root mean square (RMS) error, i.e., the partial derivative of objective function at time step s at time t. B denotes the batch size. $\mu_1, \mu_2 \in [0, 1]$ in (23) are the degree of the previous influence on the current direction to be specified by designer, which can be considered as the concept of momentum to avoid being trapped by a local minimum and speed up the learning process [23], [24]. Based on the bias-corrected estimators in (22) and (23), if the direction of current gradient $g_s(t)$ is the same as the accumulated gradient, then the gradient will be strengthened, otherwise, the gradient will be weakened. μ_1^s and μ_2^s denote the sth power of μ_1 and μ_2, respectively. $m_s(t)$ and $v_s(t)$ in (23) are

the moving average of gradient and squared gradient of $g_s(t)$ at time t, respectively. With $\tilde{v}_s(t)$ in (21), we can take the advantage of the adaptive learning rate, i.e., it should be larger at the beginning and then smaller near the minimum. The Adam learning algorithm in (21)-(23) can combine both the advantages of the above momentum and RMSProp [23], [24] and is found to be an efficient parameter-specific adaptive learning method. Due to its easy implementation and great performance, Adam learning algorithm is one of the most popular optimizer being uesd recently and will be employed to train HJIE-embedded DNN to output $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T}$ for solving $HJIE = 0$ in (19) and generating H_∞ observer gain $L^*(\hat{x}(t))$, H_∞ control $u^*(t)$ and the worst-case $\bar{v}^*(t)$ in (15), (14) and (13), respectively, for the HJIE-embedded DNN-based H_∞ observer-based output feedback reference tracking control scheme in Fig. 1.

Remark 2: The convergence of weighting parameter vector $\theta_s(t)$ of Adam learning algorithm in (21)-(23) has been proven in [23]. If the number of hidden neurons and time steps S are large enough, the updating weighting parameter vector $\theta_s(t)$ of DNN by Adam learning algorithm could converge to a globally optimal $\theta_s^*(t)$ at a linear convergence rate as $s \to \infty$ in (21)-(23).

During the off-line training process of HJIE-embedded DNN through the above Adam learning algorithm in (21)-(23) in Fig. 1, the output $(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T})_\varepsilon$ of DNN is sent to HJIE$_\varepsilon$ in (20) to calculate its error $\varepsilon(\theta_s(t))$ at the sth training step of Adam learning algorithm at time t in (21)-(23) as follows:

$$\begin{aligned} HJIE_\varepsilon &= \bar{x}^T(t)\bar{Q}\bar{x}(t) - \frac{1}{4}(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T})_\varepsilon^T \\ &\times \begin{bmatrix} \bar{G}(\tilde{x}(t),e(t),t) \\ 0 \end{bmatrix} R^{-1} \begin{bmatrix} \bar{G}(\tilde{x}(t),e(t),t) \\ 0 \end{bmatrix}^T \\ &\times (\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T})_\varepsilon + (\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T})_\varepsilon^T \\ &\times \begin{bmatrix} \bar{F}(\tilde{x}(t),e(t),t) \\ 1 \end{bmatrix} + \frac{1}{4\rho^2}(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T})_\varepsilon^T \\ &\times \begin{bmatrix} \bar{D}(\tilde{x}(t),e(t),t) & 0 \\ 0 & 0 \end{bmatrix} (\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T})_\varepsilon \\ &- \frac{1}{16\rho^2}\tilde{C}^T(\tilde{x}(t))\tilde{C}(\tilde{x}(t)) = \varepsilon(\theta_s(t)) \end{aligned} \tag{24}$$

The error $\varepsilon(\theta_s(t))$ of HJIE will be fedback to train DNN iteratively by Adam learning algorithm to expectantly output the precise $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T}$ for the H_∞ control law $u^*(t) = -\frac{1}{2}R^{-1}\bar{G}^T(\tilde{x}(t),e(t),t)\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)]^T}$, the H_∞ observer gain $L^*(\hat{x}(t)) = \frac{1}{2}\frac{\left(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial\tilde{x}(t)}\right)}{\left\|\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial\tilde{x}(t)}\right\|^2}\tilde{C}^T(\tilde{x}(t))$ and the worst-case external disturbance and measurement noise $\begin{bmatrix} v^*(t) \\ n^*(t) \end{bmatrix} = \frac{1}{2\rho^2}\begin{bmatrix} D(x(t)) & -L^*(\hat{x}(t)) \\ D_e(e(t),t) & 0 \end{bmatrix}^T \times \frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)]^T}$ as the error $\varepsilon(\theta_s(t)) \to 0$.

In the following theorem, we will prove that as the error $\varepsilon(\theta_s(t))$ in (24) approaches to 0 by the Adam learning algorithm in (21)-(23), the output $(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T})_\varepsilon$ in (24) can approach to $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T}$ in (16).

Theorem 2: If $\varepsilon(\theta_s(t))$ in (24) approaches to 0 by Adam learning algorithm in (21)-(23), then $(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T})_\varepsilon$ in (24) will approach to $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T}$ in (16). In this situation, (a) the HJIE-embedded DNN-based robust observer-based output feedback reference tracking scheme in Fig. 1 will approach to the theoretical H_∞ observer-based output feedback reference tracking control design (13)-(15) in Theorem 1. (b) If $v(t)$, $n(t)$ and $r(t) \in \mathcal{L}_2[0,\infty)$, the mean-square asymptotical estimation tracking and stability of closed loop system is also guaranteed, i.e., $\tilde{x}(t) \to 0$, $e(t) \to 0$, $u(t) \to 0$, $x(t) \to 0$ as $t_f \to \infty$ in the mean-square sense.

Proof: See Appendix B. □

From Theorem 2, it can be seen that the proposed HJIE-embedded DNN-based robust observer-based reference tracking scheme by Adam learning algorithm in Fig. 1 can approach to the theoretical H_∞ observer-based output feedback reference tracking control design in Theorem 1(a) when $\varepsilon(\theta_s(t))$ of HJIE approaches to zero after the deep learning process of Adam learning algorithm in (21)-(23). In Theorem 1(b), if the nonlinear stochastic system in (1) is with $v(t)$ and $n(t) \in \mathcal{L}_2(0,\infty)$, then the minmax H_∞ observer-based output feedback reference tracking control could achieve the asymptotical estimation and reference tracking ability, i.e., $\tilde{x}(t) \to 0$ and $e(t) \to 0$ as $t \to \infty$. Therefore, as $\varepsilon(\theta_s(t)) \to 0$, the proposed HJIE-embedded DNN-based robust observer-based output feedback control scheme in Fig. 1 can also achieve the asymptotical estimation and reference tracking ability of nonlinear stochastic system in (1) $v(t)$ and $n(t) \in \mathcal{L}_2(0,\infty)$.

Remark 3: In the off-line training phase, we input the estimation error $\tilde{x}(t)$ and the reference tracking error $e(t)$ into HJIE-embedded DNN as shown in Fig. 1. According to Theorem 2, we can train HJIE-embedded DNN to output $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T}$ to generate H_∞ control $u^*(t)$ and H_∞ observer gain $L^*(\hat{x}(t))$ if $\varepsilon(\theta_s(t))$ calculated by HJIE$_\varepsilon$ approaches to 0. However, in practical applications, we always stop the off-line training phase in Fig. 1 and shift to on-line operation phase if $|\varepsilon(\theta_s(t))| \le \delta$ for a prescribed small value δ or the number of training time steps approaches to a specified number S in (21). In this study, the number of training time steps $S = 30$ is given in the following design example.

Remark 4: In the on-line operation phase as shown in Fig. 1, based on the training weighting parameters $\theta_s(t)$ of DNN in the off-line training phase, $y(t)$ can be generated by real physical system of (1) with real external disturbance $v(t)$ and measurement noise $n(t)$ through H_∞ observer-based output feedback control $u^*(t)$. We input $\tilde{x}(t)$ and $e(t)$ into DNN to generate $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T}$ to produce $u^*(t)$ and $L^*(\hat{x}(t))$ for the robust H_∞ observer-based reference tracking control of

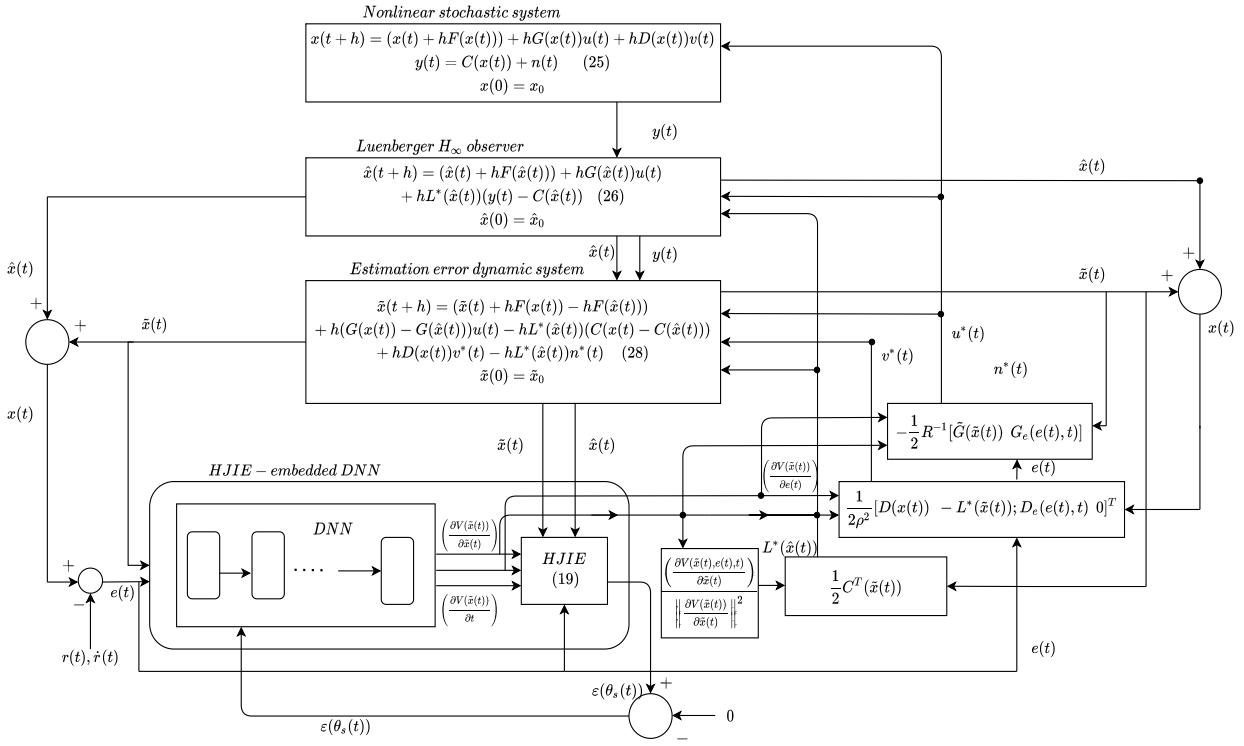

FIGURE 3. The flow chart of the HJIE-embedded DNN-based H_∞ observer-based reference tracking control of nonlinear stochastic sample-data system with external disturbance and measurement noise in (25). In the off-line operation phase, the output $y(t)$ is generated by the physical system with the worst-case external disturbance $v^*(t)$ and measurement noise $n^*(t)$ through H_∞ observer-based control input $u^*(t)$. Since $x(t)$ is unavailable, which is needed in calculating HJIE in (19), we need Luenberger observer in (26) to generate $\hat{x}(t)$ and state estimation error equation in (28) to generate $\tilde{x}(t)$ to obtain $x(t) = \hat{x}(t) + \tilde{x}(t)$ for HJIE in (19) and $e(t) = x(t) - r(t)$ as input to the trained DNN to output $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T}$ to generate $L^*(\hat{x}(t))$ in (15) for observer and estimation error dynamic and $u^*(t) = K^*(\hat{x}(t), e(t))$ in (14) for nonlinear stochastic system and estimation error equation.

nonlinear stochastic system in (1). However, if $|\varepsilon(\theta_s(t))| > \delta$ during on-line operation phase, the weighting parameters $\theta_s(t)$ of DNN can be still updated by Adam learning algorithm in (21)-(23) in the on-line operation phase without the influence on DNN-based control.

For the convenience of practical applications, the nonlinear stochastic system in (1) can be described by the following nonlinear stochastic sample-data system

$$\begin{aligned}\frac{x(t+h)-x(t)}{h} &= F(x(t)) + G(x(t))u(t)\\ &\quad +D(x(t))v(t), \quad x(0) = x_0\\ y(t) &= C(x(t)) + n(t)\end{aligned}$$

where h denotes the sampling time.

Or

$$\begin{aligned}x(t+h) &= (x(t) + hF(x(t))) + hG(x(t))u(t)\\ &\quad +hD(x(t))v(t), \quad x(0) = x_0\\ y(t) &= C(x(t)) + n(t)\end{aligned} \tag{25}$$

i.e., $F(x(t))$ is changed to $x(t) + hF(x(t))$, $G(x(t))$ is changed to $hG(x(t))$ and $D(x(t))$ is changed to $hD(x(t))$.

In this case, the observer-based output feedback control in (5) is modified as

$$\begin{aligned}\hat{x}(t+h) &= (\hat{x}(t) + hF(\hat{x}(t))) + hG(\hat{x}(t))u(t)\\ &\quad +hL(\hat{x}(t))(y(t) - C(x(t))), \quad \hat{x}(0) = \hat{x}_0\\ u(t) &= K(\hat{x}(t), e(t))\end{aligned} \tag{26}$$

Similarly, the reference tracking error dynamic in (4) and state estimation error equation in (7) are modified as

$$\begin{aligned}e(t+h) &= (e(t) + hF_e(e(t), t)) + hG_e(e(t), t)u(t)\\ &\quad +hD_e(e(t), t)v(t)\end{aligned} \tag{27}$$

and

$$\begin{aligned}\tilde{x}(t+h) &= (\tilde{x}(t) + hF(x(t)) - hF(\hat{x}(t))) + h(G(x(t))\\ &\quad -G(\hat{x}(t)))u(t)\ - hL(\hat{x}(t))(C(x(t)) - C(\hat{x}(t)))\\ &\quad +h\left[\, D(x(t)) \quad -L(\hat{x}(t))\,\right]\bar{v}(t)\end{aligned} \tag{28}$$

respectively. For the convenience of practical designs, the flow chart of HJIE-embedded DNN-based H_∞ observer-based output feedback reference tracking scheme in Fig. 1 for nonlinear stochastic continuous system in (1)-(5) is modified as Fig. 3 for nonlinear stochastic sample-data system in (25)-(28). Based on the above analyses of the nonlinear

stochastic sample-data system in (25)-(28), the flow chart of HJIE-embedded DNN-based H_∞ observer-based output feedback reference tracking control scheme is modified in Fig. 3. In the following simulation example of H_∞ observer-based reference tracking control of quadrotor UAV, for the convenience of design, the nonlinear stochastic sample-data systems (25)-(28) are employed for practical application.

Remark 5: In this study, unlike the conventional big data-driven DNN schemes, the training data $\tilde{x}(t)$ and $e(t)$ of DNN are generated by system model, observer model and estimation model, respectively. Furthermore, the theoretical result of HJIE for H_∞ observer-based reference tracking control strategy is also employed to train DNN to achieve the robust H_∞ state estimation and H_∞ reference tracking performance simultaneously, which are not easily achieved by the conventional big-data driven DNN approaches. Therefore we could save a much amount of training time and training data of HJIE-embedded DNN. Moreover, with the well-trained DNN with a large amount of initial conditions in the off-line training phase, we could also avoid the instability of the H_∞ observer-based output feedback tracking control of nonlinear stochastic system in (1) in the beginning of on-line operation phase.

Remark 6: The minmax H_∞ observer-based output feedback reference tracking control design problem in Theorem 1 can be reduced to the following optimal H_2 observer-based output feedback reference control design problem if $v(t) = 0$ and $n(t) = 0$ in (1) and $\rho^2 = \infty$ in (6), i.e., the robust H_∞ observer-based reference tracking control strategy becomes the following optimal H_2 quadratic observer-based reference tracking control strategy if $v(t) = 0, n(t) = 0$ and $\rho^2 = \infty$ [15]

$$\min_{\substack{K(\hat{x}(t),e(t))\\ L(\hat{x}(t))}} E\{\int_0^{t_f} ((x(t)-\hat{x}(t))^T Q_1(x(t)-\hat{x}(t)) + e^T(t)Q_2e(t) + u^T(t)Ru(t))dt\} \quad (29)$$

for the nonlinear system in (1). Then the H_2 optimal control gain and observer gain of the observer-based reference tracking control law in (5), which achieve the minimization in (29), are given as follows:

$$u^*(t) = K^*(\hat{x}(t), e(t)) = -\frac{1}{2}R^{-1}\bar{G}^T(\tilde{x}(t), e(t), t)\left(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)]^T}\right) \quad (30)$$

$$L^*(\hat{x}(t)) = \frac{1}{2}\frac{\left(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial\tilde{x}(t)}\right)}{\left\|\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial\tilde{x}(t)}\right\|^2}\tilde{C}^T(\tilde{x}(t)) \quad (31)$$

where $V(\tilde{x}(t), e(t), t) > 0$ with the initial condition $V(0, 0, t) = 0$ is the solution of the following HJIE,

$$HJIE = \left(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T}\right)^T \begin{bmatrix} \bar{F}(\tilde{x}(t), e(t), t) \\ 1 \end{bmatrix} + \tilde{x}^T(t)\bar{Q}\tilde{x}(t) - \frac{1}{4}\left(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T}\right)^T \times \begin{bmatrix} \bar{G}(\tilde{x}(t), e(t), t) \\ 0 \end{bmatrix} R^{-1} \begin{bmatrix} \bar{G}(\tilde{x}(t), e(t), t) \\ 0 \end{bmatrix}^T \times \left(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T}\right) = 0 \quad (32)$$

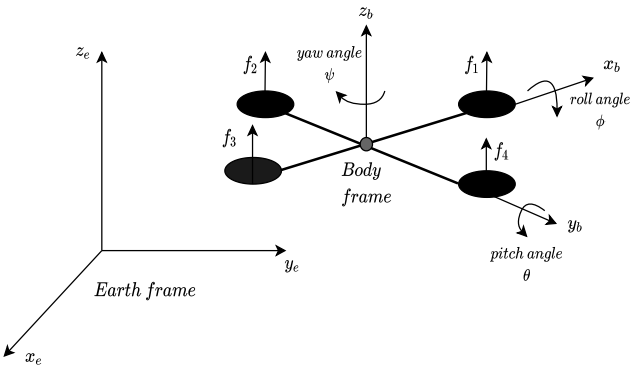

FIGURE 4. The quadrotor UAV dynamic system with the attitude and position in the inertial earth-fixed frame (x_e, y_e, z_e) and the body-fixed frame (x_b, y_b, z_b).

which is a modification of HJIE in (19) with $\rho^2 = \infty$ [11]–[13]. Therefore, with a small modification, the proposed HJIE-embedded DNN-based H_∞ observer-based control scheme in Fig. 1 and Fig. 3 can be also used to treat the optimal H_2 quadratic observer-based output feedback tracking control of nonlinear stochastic system in (1) without consideration of external disturbance and measurement noise.

Remark 7: While $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial\tilde{x}(t)} = 0$, $L^*(\hat{x}(t))$ in (31) will become singular. In this situation, the previous $L^*(\hat{x}(t-h))$ can replace $L^*(\hat{x}(t))$ to avoid the singularity of $L^*(\hat{x}(t))$.

IV. SIMULATION EXAMPLE

After the introduction of HJIE-embedded DNN-based robust H_∞ observer-based output feedback reference tracking control design for nonlinear stochastic system with external disturbance and output measurement noise in (1), an H_∞ observer-based output feedback reference tracking design example of quadrotor unmanned aerial vehicle (UAV) is given in this section to validate the reference tracking performance of the proposed scheme in comparison with robust H_∞ T-S fuzzy observer-based output feedback reference tracking control method [16]. The dynamic system of a quadrotor UAV is described by two reference frames including the inertial earth-fixed frame (x_e, y_e, z_e) and the body-fixed frame (x_b, y_b, z_b) as shown in Fig. 4.

The linear position of the mass center of the quadrotor UAV in the inertial earth-fixed frame (x_e, y_e, z_e) is denoted by $\Gamma = (x, y, z)^T \in \mathbb{R}^3$ and the orientation of the quadrotor UAV is described by Euler angles $\Theta = (\phi, \theta, \psi)^T \in \mathbb{R}^3$, i.e., roll, pitch and yaw angles, respectively, at the body-fixed frame (x_b, y_b, z_b) in Fig. 4. It is assumed that these three Eular angles within $(-\pi/2 < \phi < \pi/2)$, $(-\pi/2 < \theta < \pi/2)$ and $(-\pi < \psi < \pi)$, respectively. Therefore, the linear velocity and acceleration of quadrotor UAV are

given as $\dot{\Gamma} = (\dot{x}, \dot{y}, \dot{z})^T \in \mathbb{R}^3$ and $\ddot{\Gamma} = (\ddot{x}, \ddot{y}, \ddot{z})^T \in \mathbb{R}^3$ in the earth-fixed frame. Similarly, the angular velocity and acceleration of UAV are given as $\dot{\Theta} = (\dot{\phi}, \dot{\theta}, \dot{\psi})^T \in \mathbb{R}^3$ and $\ddot{\Theta} = (\ddot{\phi}, \ddot{\theta}, \ddot{\psi})^T \in \mathbb{R}^3$ in the body-fixed frame, respectively.

With the above definitions and notation, the dynamic equation of quadrotor UAV in Fig. 4 can be written in form of two subsystems corresponding to translational motion (referring to the position $(x(t), y(t), z(t))$ of mass center of UAV) and angular motion (referring to the attitude $(\phi(t), \theta(t), \psi(t))$ of UAV). Based on the Newton-Euler method, the stochastic quadrotor UAV system with external disturbance and output measurement noise can be described as follows [17]:

$$
\begin{aligned}
&[\dot{x}_1(t), \dot{x}_2(t), \dot{y}_1(t), \dot{y}_2(t), \dot{z}_1(t), \dot{z}_2(t) \\
&\quad , \dot{\phi}_1(t), \dot{\phi}_2(t), \dot{\theta}_1(t), \dot{\theta}_2(t), \dot{\psi}_1(t), \dot{\psi}_2(t)]^T \\
&= \begin{bmatrix}
x_2(t) \\
-\frac{K_x}{m} x_2(t) + (\cos\phi_1(t) \sin\theta_1(t) \cos\psi_1(t) \\
\quad + \sin\phi_1(t) \sin\psi_1(t)) \frac{F(t)}{m} + v_x(t) \\
y_2(t) \\
-\frac{K_y}{m} y_2(t) + (\cos\phi_1(t) \sin\theta_1(t) \sin\psi_1(t) \\
\quad - \sin\phi_1(t) \cos\psi_1(t)) \frac{F(t)}{m} + v_y(t) \\
z_2(t) \\
-\frac{K_z}{m} z_2(t) - g \\
\quad + \cos\phi_1(t) \cos\theta_1(t) \frac{F(t)}{m} + v_z(t) \\
\phi_2(t) \\
\frac{J_y - J_z}{J_x} \theta_1(t) \psi_1(t) - \frac{K_\phi}{J_x} \phi_2(t) \\
\quad + \frac{l}{J_x} \tau_\phi(t) + v_\phi(t) \\
\theta_2(t) \\
\frac{J_z - J_x}{J_y} \phi_1(t) \psi_1(t) - \frac{K_\theta}{J_y} \theta_2(t) \\
\quad + \frac{l}{J_y} \tau_\theta(t) + v_\theta(t) \\
\psi_2(t) \\
\frac{J_x - J_y}{J_z} \phi_1(t) \theta_1(t) - \frac{K_\psi}{J_z} \psi_2(t) \\
\quad + \frac{c}{J_z} \tau_\psi(t) + v_\psi(t)
\end{bmatrix}
\end{aligned} \tag{33}
$$

where $x_1(t), y_1(t), z_1(t) \in \mathbb{R}^1$ define the locations of the UAV in x-, y-, and z-axes in the Cartesian coordinate space with respect to the inertial frame, $x_2(t), y_2(t), z_2(t) \in \mathbb{R}^1$ define the velocities of the UAV in x-, y-, and z-axes in the Cartesian coordinate space with respect to the inertial frame, $\phi_1(t), \theta_1(t), \psi_1(t) \in \mathbb{R}^1$ define the attitudes of the UAV in the Euler angles with respect to the body frame and $\phi_2(t), \theta_2(t), \psi_2(t) \in \mathbb{R}^1$ define the angular velocities of the UAV in the Euler angles with respect to the body frame, respectively. $v_x(t), v_y(t), v_z(t)$ are the external disturbances of the UAV in the three translation dynamics and $v_\phi(t), v_\theta(t), v_\psi(t)$ are the external disturbances caused by unexpected rotation force in roll, pitch and yaw dynamics, respectively. $m \in \mathbb{R}^+$ is the total mass of the UAV, $J_x, J_y, J_z \in \mathbb{R}^+$ are the moments of inertia of the UAV, $g \in \mathbb{R}^+$ is the acceleration of gravity, $K_x, K_y, K_z, K_\phi, K_\theta, K_\psi \in \mathbb{R}^+$ denote the aerodynamic damping coefficients of the UAV. $F(t) \in \mathbb{R}^1$ represents the total thrust and $\tau_\phi(t), \tau_\theta(t), \tau_\psi(t) \in \mathbb{R}^1$ represent the rational torques caused by the four rotors of the UAV. $l \in \mathbb{R}^+$ denotes the distance between epicenter of the UAV and the rotor axis. $c \in \mathbb{R}^+$ denotes costant of force-to-moment factor.

In this observer-based reference tracking design example, let us denote the state vector $X(t)$ of quadrotor UAV and the desired reference target $r(t)$ to be tracked by quadrotor UAV as follows:

$$
\begin{aligned}
X(t) &\triangleq [x_1(t), x_2(t), y_1(t), y_2(t), z_1(t), z_2(t) \\
&\quad , \phi_1(t), \phi_2(t), \theta_1(t), \theta_2(t), \psi_1(t), \psi_2(t)]^T \\
&\triangleq [x(t), \dot{x}(t), y(t), \dot{y}(t), z(t), \dot{z}(t) \\
&\quad , \phi(t), \dot{\phi}(t), \theta(t), \dot{\theta}(t), \psi(t), \dot{\psi}(t)]^T \\
r(t) &\triangleq [r_1(t), r_2(t), r_3(t), r_4(t), r_5(t), r_6(t) \\
&\quad , r_7(t), r_8(t), r_9(t), r_{10}(t), r_{11}(t), r_{12}(t)]^T \\
&\triangleq [x_d(t), \dot{x}_d(t), y_d(t), \dot{y}_d(t), z_d(t), \dot{z}_d(t) \\
&\quad , \phi_d(t), \dot{\phi}_d(t), \theta_d(t), \dot{\theta}_d(t), \psi_d(t), \dot{\psi}_d(t)]^T
\end{aligned} \tag{34}
$$

Then the tracking error between the state $X(t)$ of UAV and the desired reference target $r(t)$ is defined as

$$
\begin{aligned}
e(t) &\triangleq \begin{bmatrix} e_1(t) \\ e_2(t) \\ e_3(t) \\ e_4(t) \\ e_5(t) \\ e_6(t) \\ e_7(t) \\ e_8(t) \\ e_9(t) \\ e_{10}(t) \\ e_{11}(t) \\ e_{12}(t) \end{bmatrix} \triangleq \begin{bmatrix} x_1(t) - r_1(t) \\ x_2(t) - r_2(t) \\ y_1(t) - r_3(t) \\ y_2(t) - r_4(t) \\ z_1(t) - r_5(t) \\ z_2(t) - r_6(t) \\ \phi_1(t) - r_7(t) \\ \phi_2(t) - r_8(t) \\ \theta_1(t) - r_9(t) \\ \theta_2(t) - r_{10}(t) \\ \psi_1(t) - r_{11}(t) \\ \psi_2(t) - r_{12}(t) \end{bmatrix} \triangleq \begin{bmatrix} x(t) - x_d(t) \\ \dot{x}(t) - \dot{x}_d(t) \\ y(t) - y_d(t) \\ \dot{y}(t) - \dot{y}_d(t) \\ z(t) - z_d(t) \\ \dot{z}(t) - \dot{z}_d(t) \\ \phi(t) - \phi_d(t) \\ \dot{\phi}(t) - \dot{\phi}_d(t) \\ \theta(t) - \theta_d(t) \\ \dot{\theta}(t) - \dot{\theta}_d(t) \\ \psi(t) - \psi_d(t) \\ \dot{\psi}(t) - \dot{\psi}_d(t) \end{bmatrix}
\end{aligned} \tag{35}
$$

Thus, based on (33), the stochastic quadrotor UAV system with external disturbance and measurement noise can be described as the following nonlinear stochastic system

$$
\begin{aligned}
\dot{X}(t) &= F(X(t)) + G(X(t))u(t) + D(X(t))v(t) \\
Y(t) &= C(X(t)) + n(t)
\end{aligned} \tag{36}
$$

where $X(t)$ is defined in (34) and

$$
\begin{aligned}
u(t) &= \left[F(t), \tau_\phi(t), \tau_\theta(t), \tau_\psi(t)\right]^T \\
v(t) &= \left[v_x(t), v_y(t), v_z(t), v_\phi(t), v_\theta(t), v_\psi(t)\right]^T \\
Y(t) &= [Y_1(t), Y_2(t), Y_3(t), Y_4(t), Y_5(t), Y_6(t) \\
&\quad , Y_7(t), Y_8(t), Y_9(t), Y_{10}(t), Y_{11}(t), Y_{12}(t)]^T
\end{aligned}
$$

$$F(X(t)) = \begin{bmatrix} x_2(t) \\ -\frac{K_x}{m}x_2(t) \\ y_2(t) \\ -\frac{K_x}{m}y_2(t) \\ z_2(t) \\ -\frac{K_x}{m}z_2(t) - g \\ \phi_2(t) \\ \frac{J_y - J_z}{J_x}\theta_1(t)\psi_1(t) - \frac{K_\phi}{J_x}\phi_2(t) \\ \theta_2(t) \\ \frac{J_z - J_x}{J_y}\phi_1(t)\psi_1(t) - \frac{K_\theta}{J_y}\theta_2(t) \\ \psi_2(t) \\ \frac{J_x - J_y}{J_z}\phi_1(t)\theta_1(t) - \frac{K_\psi}{J_z}\psi_2(t) \end{bmatrix}$$

$$C(X(t)) = [x_1(t), x_2(t), y_1(t), y_2(t), z_1(t), z_2(t), \phi_1(t), \phi_2(t), \theta_1(t), \theta_2(t), \psi_1(t), \psi_2(t)]^T$$

$$D(X(t)) = \begin{bmatrix} 0&0&0&0&0&0 \\ 1&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&1&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&1&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&1&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&1&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&1 \end{bmatrix}, \quad G(X(t))$$

$$= \begin{bmatrix} 0 & 0&0&0 \\ (\cos\phi_1(t)\sin\theta_1(t)\cos\psi_1(t) + \sin\phi_1(t)\sin\psi_1(t))\frac{1}{m} & 0&0&0 \\ 0 & 0&0&0 \\ (\cos\phi_1(t)\sin\theta_1(t)\sin\psi_1(t) - \sin\phi_1(t)\cos\psi_1(t))\frac{1}{m} & 0&0&0 \\ 0 & 0&0&0 \\ \cos\phi_1(t)\cos\theta_1(t)\frac{1}{m} & 0&0&0 \\ 0 & 0&0&0 \\ 0 & \frac{l}{J_x}&0&0 \\ 0 & 0&0&0 \\ 0 & 0&\frac{l}{J_y}&0 \\ 0 & 0&0&0 \\ 0 & 0&0&\frac{c}{J_z} \end{bmatrix}$$

According to Theorem 1, the following H_∞ Luenberger observer-based output feedback control is designed to achieve the H_∞ observer-based reference tracking strategy in (6) or (10)

$$\begin{aligned} \dot{\hat{X}}(t) &= F(\hat{X}(t)) + G(\hat{X}(t))u(t) \\ &\quad + L^*(\hat{X}(t))(y(t) - C(\hat{X}(t))) \\ u(t) &= K^*(\hat{X}(t), e(t)) \end{aligned} \tag{37}$$

where

$$F(\hat{X}(t)) = \begin{bmatrix} \hat{x}_2(t) \\ -\frac{K_x}{m}\hat{x}_2(t) \\ \hat{y}_2(t) \\ -\frac{K_x}{m}\hat{y}_2(t) \\ \hat{z}_2(t) \\ -\frac{K_x}{m}\hat{z}_2(t) - g \\ \hat{\phi}_2(t) \\ \frac{J_y - J_z}{J_x}\hat{\theta}_1(t)\hat{\psi}_1(t) - \frac{K_\phi}{J_x}\hat{\phi}_2(t) \\ \hat{\theta}_2(t) \\ \frac{J_z - J_x}{J_y}\hat{\phi}_1(t)\hat{\psi}_1(t) - \frac{K_\theta}{J_y}\hat{\theta}_2(t) \\ \hat{\psi}_2(t) \\ \frac{J_x - J_y}{J_z}\hat{\phi}_1(t)\hat{\theta}_1(t) - \frac{K_\psi}{J_z}\hat{\psi}_2(t) \end{bmatrix}$$

$$C(\hat{X}(t)) = [\hat{x}_1(t), \hat{x}_2(t), \hat{y}_1(t), \hat{y}_2(t), \hat{z}_1(t), \hat{z}_2(t), \hat{\phi}_1(t), \hat{\phi}_2(t), \hat{\theta}_1(t), \hat{\theta}_2(t), \hat{\psi}_1(t), \hat{\psi}_2(t)]^T$$

$$G(\hat{X}(t)) = \begin{bmatrix} 0 & 0&0&0 \\ (\cos\hat{\phi}_1(t)\sin\hat{\theta}_1(t)\cos\hat{\psi}_1(t) + \sin\hat{\phi}_1(t)\sin\hat{\psi}_1(t))\frac{1}{m} & 0&0&0 \\ 0 & 0&0&0 \\ (\cos\hat{\phi}_1(t)\sin\hat{\theta}_1(t)\sin\hat{\psi}_1(t) - \sin\hat{\phi}_1(t)\cos\hat{\psi}_1(t))\frac{1}{m} & 0&0&0 \\ 0 & 0&0&0 \\ \cos\hat{\phi}_1(t)\cos\hat{\theta}_1(t)\frac{1}{m} & 0&0&0 \\ 0 & 0&0&0 \\ 0 & \frac{l}{J_x}&0&0 \\ 0 & 0&0&0 \\ 0 & 0&\frac{l}{J_y}&0 \\ 0 & 0&0&0 \\ 0 & 0&0&\frac{c}{J_z} \end{bmatrix}$$

with the H_∞ control $u^*(t) = K^*(\hat{X}(t), e(t)) = -\frac{1}{2}R^{-1}\tilde{G}^T(\tilde{X}(t), e(t), t)(\frac{\partial V(\tilde{X}(t), e(t), t)}{\partial [\tilde{X}^T(t)\ e^T(t)]^T})$ in (14) and H_∞ observer gain $L^*(\hat{X}(t)) = \frac{1}{2}\frac{(\frac{\partial V(\tilde{X}(t), e(t), t)}{\partial \tilde{X}(t)})}{\left\|\frac{\partial V(\tilde{X}(t), e(t), t)}{\partial \tilde{X}(t)}\right\|^2}\tilde{C}^T(\tilde{X}(t))$ in (15), where the estimation error $\tilde{X}(t)$ and reference tracking error $e(t)$ of the quadrotor UAV are generated, respectively, by

$$\begin{aligned} \dot{\tilde{X}}(t) &= \tilde{F}(\tilde{X}(t)) + \tilde{G}(\tilde{X}(t))u(t) - L^*(\hat{X}(t))\tilde{C}(\tilde{X}(t)) \\ &\quad + \left[\, D(X(t)) \ -L^*(\hat{X}(t)) \,\right]\bar{v}(t) \end{aligned} \tag{38}$$

where

$$\tilde{F}(\tilde{X}(t)) = F(X(t)) - F(\hat{X}(t))$$

$$= \begin{bmatrix} \tilde{x}_2(t) \\ -\frac{K_x}{m}\tilde{x}_2(t) \\ \tilde{y}_2(t) \\ -\frac{K_x}{m}\tilde{y}_2(t) \\ \tilde{z}_2(t) \\ -\frac{K_x}{m}\tilde{z}_2(t) \\ \tilde{\phi}_2(t) \\ \frac{J_y-J_z}{J_x}(\theta_1(t)\psi_1(t) - \hat{\theta}_1(t)\hat{\psi}_1(t)) - \frac{K_\phi}{J_x}\tilde{\phi}_2(t) \\ \tilde{\theta}_2(t) \\ \frac{J_z-J_x}{J_y}(\phi_1(t)\psi_1(t) - \hat{\phi}_1(t)\hat{\psi}_1(t)) - \frac{K_\theta}{J_y}\tilde{\theta}_2(t) \\ \tilde{\psi}_2(t) \\ \frac{J_x-J_y}{J_z}(\phi_1(t)\theta_1(t) - \hat{\phi}_1(t)\hat{\theta}_1(t)) - \frac{K_\psi}{J_z}\tilde{\psi}_2(t) \end{bmatrix}$$

$$\tilde{C}(\tilde{X}(t)) = C(X(t)) - C(\hat{X}(t)) = \tilde{X}(t)$$

$$\tilde{G}(\tilde{X}(t)) = G(X(t)) - G(\hat{X}(t))$$

$$= \begin{bmatrix} 0 & 0\ 0\ 0 \\ \begin{matrix}(\cos\phi_1(t)\sin\theta_1(t)\cos\psi_1(t) \\ -\cos\hat{\phi}_1(t)\sin\hat{\theta}_1(t)\cos\hat{\psi}_1(t) \\ +\sin\phi_1(t)\sin\psi_1(t) \\ -\sin\hat{\phi}_1(t)\sin\hat{\psi}_1(t))\frac{1}{m}\end{matrix} & 0\ 0\ 0 \\ 0 & 0\ 0\ 0 \\ \begin{matrix}(\cos\phi_1(t)\sin\theta_1(t)\sin\psi_1(t) \\ -\cos\hat{\phi}_1(t)\sin\hat{\theta}_1(t)\sin\hat{\psi}_1(t) \\ -\sin\phi_1(t)\cos\psi_1(t) \\ +\sin\hat{\phi}_1(t)\cos\hat{\psi}_1(t))\frac{1}{m}\end{matrix} & 0\ 0\ 0 \\ 0 & 0\ 0\ 0 \\ \begin{matrix}(\cos\phi_1(t)\cos\theta_1(t) \\ -\cos\hat{\phi}_1(t)\cos\hat{\theta}_1(t))\frac{1}{m}\end{matrix} & 0\ 0\ 0 \\ 0 & 0\ 0\ 0 \\ 0 & 0\ 0\ 0 \\ 0 & 0\ 0\ 0 \\ 0 & 0\ 0\ 0 \\ 0 & 0\ 0\ 0 \\ 0 & 0\ 0\ 0 \end{bmatrix}$$

and

$$\dot{e}(t) = F_e(e(t), t) + G_e(e(t), t)u^*(t) + D_e(e(t), t)v(t) \quad (39)$$

where

$$F_e(e(t), t) = F(e(t) + r(t)) - \dot{r}(t)$$

$$= \begin{bmatrix} e_2(t) \\ -\frac{K_x}{m}(e_2(t) + r_2(t)) - \dot{r}_2(t) \\ e_4(t) \\ -\frac{K_y}{m}(e_4(t) + r_4(t)) - \dot{r}_4(t) \\ e_6(t) \\ -\frac{K_z}{m}(e_6(t) + r_6(t)) - g - \dot{r}_6(t) \\ e_8(t) \\ \frac{J_y-J_z}{J_x}(e_9(t) + r_9(t))(e_{11}(t) + r_{11}(t)) \\ -\frac{K_\phi}{J_x}(e_8(t) + r_8(t)) - \dot{r}_8(t) \\ e_{10}(t) \\ \frac{J_z-J_x}{J_y}(e_7(t) + r_7(t))(e_{11}(t) + r_{11}(t)) \\ -\frac{K_\theta}{J_y}(e_{10}(t) + r_{10}(t)) - \dot{r}_{10}(t) \\ e_{12}(t) \\ \frac{J_x-J_y}{J_z}(e_7(t) + r_7(t))(e_9(t) + r_9(t)) \\ -\frac{K_\psi}{J_z}(e_{12}(t) + r_{12}(t)) - \dot{r}_{12}(t) \end{bmatrix}$$

$$D_e(e(t), t) = D(X(t))$$

$$G_e(e(t), t) = G(e(t) + r(t))$$

$$= \begin{bmatrix} 0 & 0\ 0\ 0 \\ \begin{matrix}(\cos(e_7(t) + r_7(t))\sin(e_9(t) \\ +r_9(t))\cos(e_{11}(t) + r_{11}(t)) \\ +\sin(e_7(t) + r_7(t)) \\ \times\sin(e_{11}(t) + r_{11}(t)))\frac{1}{m}\end{matrix} & 0\ 0\ 0 \\ 0 & 0\ 0\ 0 \\ \begin{matrix}(\cos(e_7(t) + r_7(t))\sin(e_9(t) \\ +r_9(t))\sin(e_{11}(t) + r_{11}(t)) \\ -\sin(e_7(t) + r_7(t)) \\ \times\cos(e_{11}(t) + r_{11}(t)))\frac{1}{m}\end{matrix} & 0\ 0\ 0 \\ 0 & 0\ 0\ 0 \\ \begin{matrix}\cos(e_7(t) + r_7(t)) \\ \times\cos(e_9(t) + r_9(t))\frac{1}{m}\end{matrix} & 0\ 0\ 0 \\ 0 & 0\ 0\ 0 \\ 0 & \frac{l}{J_x}\ 0\ 0 \\ 0 & 0\ 0\ 0 \\ 0 & 0\ \frac{l}{J_y}\ 0 \\ 0 & 0\ 0\ 0 \\ 0 & 0\ 0\ \frac{c}{J_z} \end{bmatrix}$$

Since we have only 4 controls, i.e., $u(t) = \left[F(t), \tau_\phi(t), \tau_\theta(t), \tau_\psi(t)\right]^T$, it is not easy to arbitrarily control 6 motions, i.e., the position $(x(t), y(t), z(t))$ in the translational motion and the attitude $(\phi(t), \theta(t), \psi(t))$ in the angular motion, independently due to the underactuated property of the quadrotor UAV [17]. Therefore, we can only control 4 motions directly and other 2 motions must be dependent on other motions. In this situation, only the desired $x_d(t), y_d(t), z_d(t)$ and the desired yaw $\psi_d(t)$ are arbitrarily selected to construct as the reference guidance system to control $x(t), y(t), z(t)$ and $\psi(t)$ to track $x_d(t), y_d(t), z_d(t)$ and $\psi_d(t)$ directly. The trajectory of yaw $\psi(t)$ is specified by the designer and is set to track $\psi_d(t) = 0$ in this example. Further, the desired roll reference $\phi_d(t)$ and pitch reference $\theta_d(t)$ are computed accordingly as follows [17]:

$$\phi_d(t) = \sin^{-1}\left(\frac{m}{F(t)}(u_x(t)\sin\psi_d(t) - u_y(t)\cos\psi_d(t))\right)$$

$$\theta_d(t) = \tan^{-1}\left(\frac{u_x(t)\cos\psi_d(t) + u_y(t)\cos\psi_d(t)}{u_z(t) + g}\right) \quad (40)$$

where the total thrust $F(t)$ and the virtual inputs $u_x(t), u_y(t), u_z(t)$ are defined as follows:

$$\begin{aligned} F(t) &= m\sqrt{u_x^2(t) + u_y^2(t) + (u_z(t) + g)^2} \\ u_x(t) &= \ddot{x}_d(t) \\ u_y(t) &= \ddot{y}_d(t) \\ u_z(t) &= \ddot{z}_d(t) \end{aligned} \quad (41)$$

where $\ddot{x}_d(t)$, $\ddot{y}_d(t)$, $\ddot{z}_d(t)$ are the double derivative of $x_d(t)$, $y_d(t)$, $z_d(t)$, respectively, i.e., in this design example of quadrotor UAV reference tracking control, we could specify the desired trajectories of $x_d(t), y_d(t), z_d(t)$ and $\psi_d(t)$ but the desired trajectories of $\phi_d(t)$ and $\theta_d(t)$ are consequentially generated by (40) and (41).

In this design example, the parameters of quadrotor UAV in (33) are given as follows: $K_x = K_y = K_z = 0.01Ns^2/rad$, $J_x = J_y = J_z = 0.1Ns^2/rad$, $m = 2kg$, $l = 1.2m$, $K_\phi = K_\theta = K_\psi = 0.01Ns^2/rad$, $c = 1$, $g = 9.8m^2/s$. The weighting matrices $Q_1 = 0.01diag\{2I_6, 5I_6\}$, $Q_2 = 0.01I_{12}$ and $R = 10^{-4}I_4$ are given in the H_∞ observer-based reference tracking control strategy in (6). The sampling time h is $0.01s$ and the terminal time $t_f = 30s$. The random external disturbances are supposed Gaussian noises with the probability distribution functions as follows: $v_x(t) \doteqdot N(0, 0.1)$, $v_y(t) \doteqdot N(0, 0.1)$, $v_z(t) \doteqdot N(0, 0.1)$, $v_\phi(t) \doteqdot N(0, 0.01)$, $v_\theta(t) \doteqdot N(0, 0.01)$ and $v_\psi(t) \doteqdot N(0, 0.01)$ where $N(0, \sigma)$ denotes the Gaussian distribution with 0 mean and σ standard deviation. The distribution function of measurement noise $n(t) \doteqdot N(0, 0.1I_{12})$. The desired reference position and yaw angle trajectories are given as $x_d(t) = 10\sin(0.5t)$, $y_d(t) = 10\cos(0.5t)$, $z_d(t) = t$ and $\psi_d(t) = 0$, respectively, and $\phi_d(t)$, $\theta(t)$ are given in (40). This is a circle with a radius of 10 meters in $x - y$ axis, rising at a constant velocity on the $z-$axis.

According to Theorem 1, the minmax H_∞ observer-based reference tracking control design needs to solve the following HJIE for H_∞ control gain $K^*(\hat{X}(t), e(t))$ in (14) and H_∞ observer gain $L^*(\hat{X}(t))$ in (15) for the observer-based output feedback control in (37) to achieve the robust H_∞ observer-based reference tracking control strategy in (6) with $\rho^2 = 2$.

$$\begin{aligned} HJIE = {} & \bar{X}^T(t)\bar{Q}\bar{X}(t) - \frac{1}{4}\Big(\frac{\partial V(\tilde{X}(t), e(t), t)}{\partial[\tilde{X}^T(t)\ e^T(t)\ t]^T}\Big)^T \\ & \times \begin{bmatrix} \bar{G}(\tilde{X}(t), e(t), t) \\ 0 \end{bmatrix} R^{-1} \\ & \times \begin{bmatrix} \bar{G}(\tilde{X}(t), e(t), t) \\ 0 \end{bmatrix}^T \Big(\frac{\partial V(\tilde{X}(t), e(t), t)}{\partial[\tilde{X}^T(t)\ e^T(t)\ t]^T}\Big) \\ & + \Big(\frac{\partial V(\tilde{X}(t), e(t), t)}{\partial[\tilde{X}^T(t)\ e^T(t)\ t]^T}\Big)^T \begin{bmatrix} \bar{F}(\tilde{X}(t), e(t), t) \\ 1 \end{bmatrix} + \frac{1}{4\rho^2} \\ & \times \Big(\frac{\partial V(\tilde{X}(t), e(t), t)}{\partial[\tilde{X}^T(t)\ e^T(t)\ t]^T}\Big)^T \begin{bmatrix} \bar{D}(\tilde{X}(t), e(t), t) & 0 \\ 0 & 0 \end{bmatrix} \\ & \times \Big(\frac{\partial V(\tilde{X}(t), e(t), t)}{\partial[\tilde{X}^T(t)\ e^T(t)\ t]^T}\Big) - \frac{1}{16\rho^2}\tilde{C}^T(\tilde{X}(t))\tilde{C}(\tilde{X}(t)) \\ = {} & 0 \end{aligned} \tag{42}$$

where the elements in augmented matrices $\bar{X}(t)$, $\tilde{C}(\tilde{X}(t))$, $\bar{F}(\tilde{X}(t), e(t), t)$, $\bar{G}(\tilde{X}(t), e(t), t)$, $\bar{D}(\tilde{X}(t), e(t), t)$ of the estimation error and reference tracking error of quadrotor UAV are defined in (38) and (39), respectively.

Since it is very difficult to solve $\frac{\partial V(\tilde{X}(t), e(t), t)}{\partial[\tilde{X}^T(t)\ e^T(t)\ t]^T}$ from (42) analytically or numerically for H_∞ control $u^*(t) = -\frac{1}{2}R^{-1}\bar{G}^T(\tilde{X}(t), e(t), t)\Big(\frac{\partial V(\tilde{X}(t), e(t), t)}{\partial[\tilde{X}^T(t)\ e^T(t)]^T}\Big)$ and H_∞ observer gain $L^*(\hat{X}(t)) = \frac{1}{2}\frac{\Big(\frac{\partial V(\tilde{X}(t), e(t), t)}{\partial \tilde{X}(t)}\Big)}{\Big\|\frac{\partial V(\tilde{X}(t), e(t), t)}{\partial \tilde{X}(t)}\Big\|^2}\tilde{C}^T(\tilde{X}(t))$ in Theorem 1 for the H_∞ observer-based output feedback controller in (5) to achieve the H_∞ reference tracking of quadrotor UAV. Therefore, the proposed HJIE-embedded DNN-based H_∞ observer-based reference tracking scheme based on stochastic sample-data systems in (25)-(28) with sampling time $h = 0.01$ in Fig. 3 is employed to solve $\frac{\partial V(\tilde{X}(t), e(t), t)}{\partial[\tilde{X}^T(t)\ e^T(t)\ t]^T}$ from HJIE in (42) for the H_∞ observer-based output feedback reference tracking control of quadrotor UAV. The architecture of DNN used in this UAV design example consists of input layer with

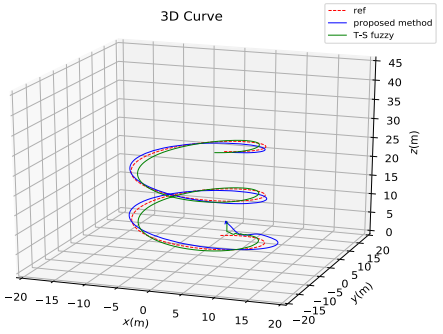

FIGURE 5. 3-D graph of the desired trajectory and the UAV flighting trajectories by the proposed H_∞ HJIE-embedded DNN-based observer-based tracking control scheme and the H_∞ observer-based T-S fuzzy tracking control scheme [16].

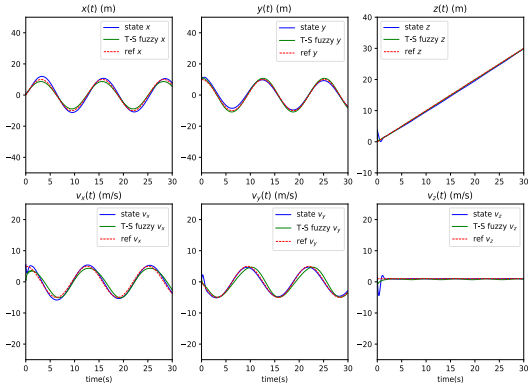

FIGURE 6. The desired references and their corresponding position trajectories by the proposed H_∞ HJIE-embedded DNN-based observer-based tracking control scheme and the H_∞ observer-based T-S fuzzy tracking control scheme [16].

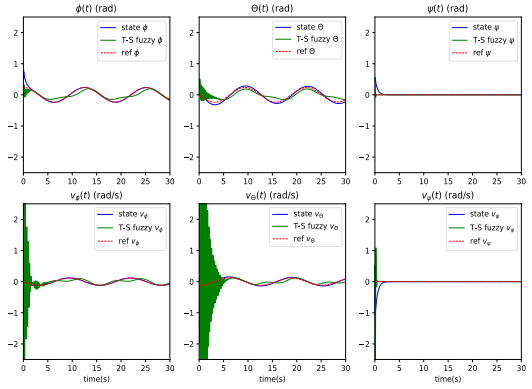

FIGURE 7. The desired references and their corresponding attitude trajectories by the proposed H_∞ HJIE-embedded DNN-based observer-based tracking control scheme and the H_∞ observer-based T-S fuzzy tracking control scheme [16].

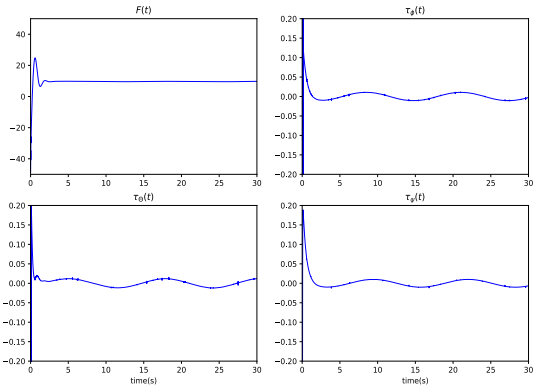

FIGURE 8. Control inputs of the UAV by the proposed H_∞ HJIE-embedded DNN-based observer-based controller.

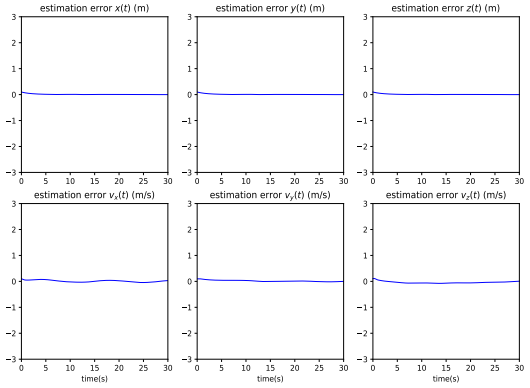

FIGURE 9. The estimation errors of the position by the proposed H_∞ HJIE-embedded DNN-based observer-based tracking control scheme.

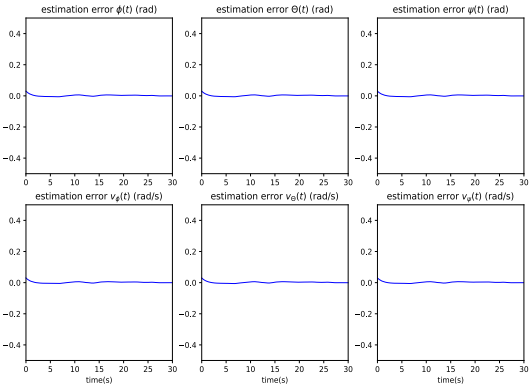

FIGURE 10. The estimation errors of the attitude by the proposed H_∞ HJIE-embedded DNN-based observer-based tracking control scheme.

input $\tilde{X}(t)$, $e(t)$ and output layer with output $\frac{\partial V(\tilde{X}(t),e(t),t)}{\partial[\tilde{X}^T(t)\ e^T(t)\ t]^T}$, 4 hidden layers, an HJIE layer with input $\frac{\partial V(\tilde{X}(t),e(t),t)}{\partial[\tilde{X}^T(t)\ e^T(t)\ t]^T}$ and output $\varepsilon(\theta(t))$ as shown in Fig. 2. There are 512, 256, 64 and 25 hidden neurons in each hidden layer sequentially. The parameters of Adam learning algorithm in (21)-(23) are set as $l = 10^{-3}$, $\mu_1 = 0.9$, $\mu_2 = 0.999$ and $\zeta = 10^{-7}$. Training steps and batch size are set as $S = 30$ and $B = 800$, respectively. In the off-line training phase in Fig. 3, we randomly select 20000 initial tracking errors $e(0)$ and estimation errors $\tilde{X}(0)$ around the origin $X(0)$. The initial state of the UAV and its state estimation are assumed as $X(0) = [0.54, 5.55, 10.57, 0.57, 3, 1.5, 0.52, 0.52, 0.51, 0.59, 0.55, 0.52]^T$ and $\hat{X}(0) = [0.44, 5.45, 10.47, 0.47, 2.9, 1.4, 0.49, 0.49, 0.48, 0.56, 0.52, 0.49]^T$.

The simulation results of the proposed H_∞ HJIE-embedded DNN-based observer-based reference tracking control scheme of quadrotor UAV are shown in Fig. 5-10. The 3-D flighting trajectory of the UAV and its desired trajectory are shown in Fig. 5. In Figs. 6-7, we see that the position and attitude trajectories can track the corresponding desired trajectories by the HJIE-embedded DNN-based observer-based control scheme. The control inputs of the UAV are constructed by the proposed method as shown in Fig. 8. In Figs. 9-10, the state estimation can be efficiently achieved by the proposed HJIE-embedded DNN-based observer-based control method with external disturbance and measurement noise. The simulation example of sinusoidal external disturbance $v_x(t) \doteqdot 0.1\sin(0.5t)$, $v_y(t) \doteqdot 0.1\sin(0.5t)$, $v_z(t) \doteqdot 0.1\cos(0.5t)$, $v_\phi(t) \doteqdot 0.01\cos(0.5t)$, $v_\theta(t) \doteqdot 0.01\sin(0.5t)$ and $v_\psi(t) \doteqdot 0.01\sin(0.5t)$ is also shown in Fig. 13. In the transient response, the UAV has fluctuation in the proposed method due to the initial value of control inputs. In the real application with DNN method, at the beginning of off-line training phase, we need to randomly select the initial training data near the state estimation error $\tilde{x}(t)$ and reference tracking error $e(t)$. This will significantly affect the training performance. If the domain of initial conditions is limited by the random selection in the off-line training phase, the state $x(t) = \hat{x}(t) + \tilde{x}(t)$ and $e(t) = x(t) + r(t)$ may be far from the training data during the on-line operation phase. Hence, it will limit the domain of $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T_\varepsilon}$ to approach to the real $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T}$ of HJIE which will cause the fluctuation in the transient response. In this situation, the error $\varepsilon(\theta_k(t))$ will be larger than a small prescribed δ, (i.e., $|\varepsilon(\theta_k(t))| > \delta$) so that we need to start on Adam learning algorithm again for updating weighting parameters of DNN. The computational complexity of H_∞ HJIE-embedded DNN-based observer-based reference tracking control scheme can be approximately calculated as $O(Ln)$, where L is denoted as the number of layers in the DNN and n is denoted as the dimension of the state estimation error $\tilde{x}(t)$ and reference tracking error $e(t)$. The real H_∞ observer-based output feedback reference tracking control performance in (10) is also calculated as follows:

$$\frac{\int_0^{30}(\bar{x}^T(t)\bar{Q}\bar{x}(t) + u^T(t) \times Ru(t))dt - V(\tilde{x}(0), e(0), 0)}{\int_0^{30}\bar{v}^T(t)\bar{v}(t)dt} \approx 1.81 \leq 2$$

In comparison with the proposed method, the H_∞ observer-based T-S fuzzy reference tracking control design in [16] is carried out and the results are shown

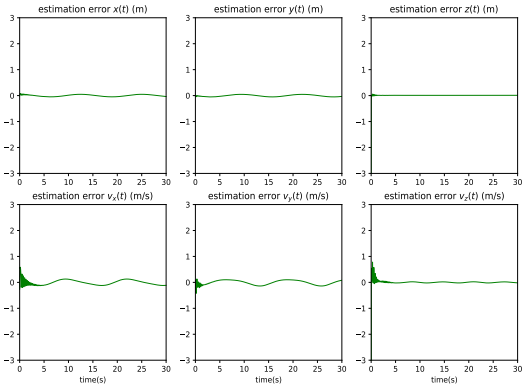

FIGURE 11. The estimation errors of the position by the H_∞ observer-based T-S fuzzy tracking control scheme [16].

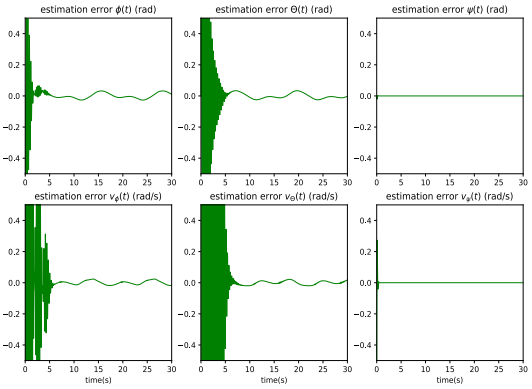

FIGURE 12. The estimation errors of the attitude by the H_∞ observer-based T-S fuzzy tracking control scheme [16].

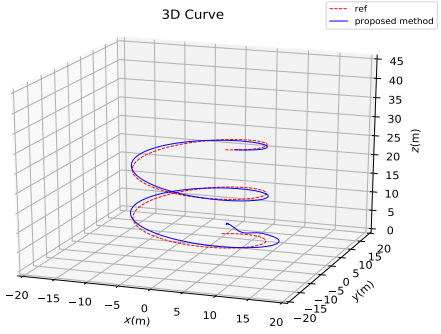

FIGURE 13. 3-D graph of the desired trajectory and the UAV flighting trajectories by the proposed H_∞ HJIE-embedded DNN-based observer-based tracking control scheme with sinusoidal external disturbance.

in Figs. 5-7, 11-12 when compared with the proposed H_∞ HJIE-embedded DNN-based observer-based reference tracking control scheme. In the transient response, the simulation results of H_∞ observer-based T-S fuzzy reference tracking control have large fluctuations in the position and attitude except for the $x(t)$, $y(t)$ and $z(t)$. Due to the limitaion of the number in fuzzy rules, the H_∞ observer-based T-S fuzzy reference tracking control design is only employed to track a small size of reference trajectory. The way to improve the tracking performance is to increase the number of fuzzy rules. However, the observer-based T-S fuzzy control law

$$\dot{\hat{x}}(t) = \sum_{i=1}^{125}\sum_{j=1}^{125} h_i(z(t))h_j(z(t))(A_i\hat{x}(t)+B_iu(t)+L_i(y-C_j\hat{x}(t)))$$

and $u(t) = \sum_{i=1}^{125} h_i(z(t))K_i\hat{x}(t)$ with 125 fuzzy rules is used in this simulation. We need to solve 125^2 BMIs for fuzzy controller and fuzzy observer in the H_∞ observer-based T-S fuzzy control design. If we want to increase the fuzzy rules to increase the performance, more computations will be required. In addition, the robust H_∞ observer-based T-S fuzzy controller needs to compute the above observer and control law at every time instant. The real tracking control performance of the H_∞ observer-based T-S fuzzy reference tracking control design is also calculated as $\rho^* \approx 7.61$ which is poor than the proposed method because of the worse tracking results in the transient response. The computational complexity of H_∞ observer-based T-S fuzzy reference tracking scheme is calculated as $O(L^{'2}n)$, where $L^{'}$ denotes as the number of the local systems and n is the dimension of the state estimation error $\tilde{x}(t)$ and reference tracking error $e(t)$.

Remark 8: In the field of observer-based tracking control designs at present, there exists no study to use the conventional big data-driven DNN method to deal with this nonlinear observer-based control design problem due to the unavailable empirical data for training DNN. In general, the conventional DNN is based on a very large amount of input/output empirical data pairs to train the weighting parameters of hidden layers in DNN by optimizer from the big data perspective. However, in the nonlinear H_∞ observer-based reference tracking control design problem, $x(t)$ is unavailable and needs to be obtained from $x(t) = \hat{x}(t)+\tilde{x}(t)$. Therefore, we can not compare with the conventional big data-driven DNN method in nonlinear H_∞ observer-based reference tracking control design. Instead, we show the H_∞ observer-based T-S fuzzy reference tracking control design to compare the tracking control performance with the proposed H_∞ HJIE-embedded DNN-based observer-based reference tracking control scheme.

V. CONCLUSION

In this study, a novel HJIE-embedded DNN-based H_∞ observer-based control scheme is proposed to directly solve $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T}$ of the nonlinear partial differential HJIE of robust H_∞ observer-based reference tracking control design problem of nonlinear stochastic system under external disturbance and measurement noise. In order to overcome the unavailable external disturbance $v(t)$ and measurement noise $n(t)$ for system model to generate training data and unavailable $x(t)$ and $e(t) = x(t) - r(t)$ in HJIE, we used $v^*(t)$ and $n^*(t)$ to replace $v(t)$ and $n(t)$ without influence on the H_∞ estimation and control, and employed estimation dynamic to generate $\hat{x}(t)$ and estimation error model to generate $\tilde{x}(t)$ so that we could obtain $x(t) = \hat{x}(t) + \tilde{x}(t)$ and tracking

error $e(t) = x(t) - r(t)$ so that we could calculate the error $HJIE_\varepsilon = \varepsilon(\theta(t))$ in (20) for Adam learning algorithm to train DNN to solve $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ e^T(t)\ t]^T}$ of HJIE for the design of H_∞ observer-based reference tracking control design. We have proven that if the approximation error of HJIE approaches to 0 by DNN through Adam learning algorithm, the proposed DNN-based estimator-based reference tracking control scheme could achieve the H_∞ estimator-based reference tracking control strategy unlike the conventional big data-driven DNN only for classification and recognition. Based on system modeling and theoretical result of HJIE for nonlinear H_∞ estimator-based reference tracking control of nonlinear stochastic system, the proposed DNN-based scheme could efficiently achieve the nonlinear H_∞ estimation and reference tracking control design simultaneously with much saving of training data and time. Finally, a design example of H_∞ observer-based reference tracking control quadrotor UAV system is also given to illustrate the design procedure and validate the effectiveness of the proposed method in comparison with the conventional robust H_∞ T-S fuzzy observer-based tracking control of nonlinear dynamic system.

APPENDIX A: PROOF OF THEOREM 1

By the indirect method, the minmax H_∞ observer-based reference tracking strategy in (10) is transformed to the equivalent constrained minmax Nash quadratic game strategy in (11). Based on two-step method, in the first step, we need to solve the minmax Nash quadratic game problem in (12). In the second step, we need to guarantee $J \le E\{V(\tilde{x}(0), e(0), 0)\}$. Therefore, Theorem 1 will be proven as follows:

(a)(i) Step 1:

From (12), we get

$$\begin{aligned} J = & \min_{\substack{u(t)=K(\hat{x}(t),e(t))\\ L(\hat{x}(t))}} \max_{\bar{v}(t)\in\mathcal{L}_2[0,t_f]} E\{\int_0^{t_f} (\bar{x}^T(t)\bar{Q}\bar{x}(t) \\ & +u^T(t)Ru(t) - \rho^2\bar{v}^T(t)\bar{v}(t))dt\} \\ = & \min_{\substack{u(t)=K(\hat{x}(t),e(t))\\ L(\hat{x}(t))}} \max_{\bar{v}(t)\in\mathcal{L}_2[0,t_f]} E\{-V(\tilde{x}(t_f), e(t_f), t_f) \\ & +V(\tilde{x}(0), e(0), 0) + \int_0^{t_f} (\bar{x}^T(t)\bar{Q}\bar{x}(t) \\ & +u^T(t)Ru(t) - \rho^2\bar{v}^T(t)\bar{v}(t) \\ & +\frac{d}{dt}V(\tilde{x}(t), e(t), t))dt\} \end{aligned} \tag{A1}$$

By the fact $V(\tilde{x}(t), e(t), t)$ is the Lyapunov function of the augmented time-varying nonlinear stochastic error system in (9), by chain rule, we get

$$\begin{aligned} \frac{dV(\tilde{x}(t), e(t), t)}{dt} = & \frac{\partial V(\tilde{x}(t), e(t), t)}{\partial t} + (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial[\tilde{x}^T(t)\ e^T(t)]^T})^T \\ & \times \left(\begin{bmatrix} \tilde{F}(\tilde{x}(t)) - L(\hat{x}(t))\tilde{C}(\tilde{x}(t)) \\ F_e(e(t), t) \end{bmatrix}\right. \\ & + \begin{bmatrix} \tilde{G}(\tilde{x}(t)) \\ G_e(e(t), t) \end{bmatrix} u(t) \\ & \left. + \begin{bmatrix} D(x(t)) & -L(\hat{x}(t)) \\ D_e(e(t), t) & 0 \end{bmatrix} \bar{v}(t)\right) \end{aligned} \tag{A2}$$

Substituting $\bar{G}(\tilde{x}(t), e(t), t) = \begin{bmatrix} \tilde{G}(\tilde{x}(t)) \\ G_e(e(t), t) \end{bmatrix}$ and (A2) into (A1), we get

$$\begin{aligned} J = & \min_{\substack{u(t)=K(\hat{x}(t),e(t))\\ L(\hat{x}(t))}} \max_{\bar{v}(t)\in\mathcal{L}_2[0,t_f]} E\{-V(\tilde{x}(t_f), e(t_f), t_f) \\ & +V(\tilde{x}(0), e(0), 0) + \int_0^{t_f} (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial t} + \bar{x}^T(t)\bar{Q}\bar{x}(t) \\ & +u^T(t)Ru(t) + (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial[\tilde{x}^T(t)\ e^T(t)]^T})^T \\ & \times \begin{bmatrix} \tilde{F}(\tilde{x}(t)) - L(\hat{x}(t))\tilde{C}(\tilde{x}(t)) \\ F_e(e(t), t) \end{bmatrix} + (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial[\tilde{x}^T(t)\ e^T(t)]^T})^T \\ & \times \bar{G}(\tilde{x}(t), e(t), t)u(t) + (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial[\tilde{x}^T(t)\ e^T(t)]^T})^T \\ & \times \begin{bmatrix} D(x(t)) & -L(\hat{x}(t)) \\ D_e(e(t), t) & 0 \end{bmatrix} \bar{v}(t) - \rho^2\bar{v}^T(t)\bar{v}(t))dt\} \end{aligned} \tag{A3}$$

By using the completing square method, we immediately have

$$\begin{aligned} J = & \min_{\substack{u(t)=K(\hat{x}(t),e(t))\\ L(\hat{x}(t))}} \max_{\bar{v}(t)\in\mathcal{L}_2[0,t_f]} E\{-V(\tilde{x}(t_f), e(t_f), t_f) \\ & +V(\tilde{x}(0), e(0), 0) + \int_0^{t_f} (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial t} + \bar{x}^T(t)\bar{Q}\bar{x}(t) \\ & -\frac{1}{4}(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial[\tilde{x}^T(t)\ e^T(t)]^T})^T \bar{G}(\tilde{x}(t), e(t), t)R^{-1} \\ & \times \bar{G}^T(\tilde{x}(t), e(t), t)(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial[\tilde{x}^T(t)\ e^T(t)]^T}) \\ & +(\frac{1}{2}\bar{G}(\tilde{x}(t), e(t), t)^T(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial[\tilde{x}^T(t)\ e^T(t)]^T}) + Ru(t))^T R^{-1} \\ & \times(\frac{1}{2}\bar{G}(\tilde{x}(t), e(t), t)^T(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial[\tilde{x}^T(t)\ e^T(t)]^T}) + Ru(t)) \\ & +(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial[\tilde{x}^T(t)\ e^T(t)]^T})^T \begin{bmatrix} \tilde{F}(\tilde{x}(t)) - L(\hat{x}(t))\tilde{C}(\tilde{x}(t)) \\ F_e(e(t), t) \end{bmatrix} \\ & +\frac{1}{4\rho^2}(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial[\tilde{x}^T(t)\ e^T(t)]^T})^T \begin{bmatrix} D(x(t)) & -L(\hat{x}(t)) \\ D_e(e(t), t) & 0 \end{bmatrix} \\ & \times \begin{bmatrix} D(x(t)) & -L(\hat{x}(t)) \\ D_e(e(t), t) & 0 \end{bmatrix}^T (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial[\tilde{x}^T(t)\ e^T(t)]^T}) \\ & -(\frac{1}{2\rho}\begin{bmatrix} D(x(t)) & -L(\hat{x}(t)) \\ D_e(e(t), t) & 0 \end{bmatrix}^T (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial[\tilde{x}^T(t)\ e^T(t)]^T}) \\ & -\rho\bar{v}(t))^T(\frac{1}{2\rho}\begin{bmatrix} D(x(t)) & -L(\hat{x}(t)) \\ D_e(e(t), t) & 0 \end{bmatrix}^T \\ & \times(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial[\tilde{x}^T(t)\ e^T(t)]^T}) - \rho\bar{v}(t)))dt\} \end{aligned} \tag{A4}$$

From (A4), it is seen that the worst-case $\begin{bmatrix} v^*(t) \\ n^*(t) \end{bmatrix}$ in (13) and the optimal $u^*(t) = K(\hat{x}(t), e(t))$ in (14)

make the involved quadratic terms be zero achieve $\min_{u(t)=K(\hat{x}(t),e(t))} \max \begin{bmatrix} v(t) \\ n(t) \end{bmatrix}$. Then we get

$$\begin{aligned} J = & \min_{L(\hat{x}(t))} E\{-V(\tilde{x}(t_f), e(t_f), t_f) + V(\tilde{x}(0), e(0), 0) \\ & + \int_0^{t_f} (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial t} + \bar{x}^T(t)\bar{Q}\bar{x}(t) \\ & + (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial [\tilde{x}^T(t) \;\; e^T(t)]^T})^T \begin{bmatrix} \tilde{F}(\tilde{x}(t)) - L(\hat{x}(t))\tilde{C}(\tilde{x}(t)) \\ F_e(e(t), t) \end{bmatrix} \\ & - \frac{1}{4}(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial [\tilde{x}^T(t) \;\; e^T(t)]^T})^T \bar{G}(\tilde{x}(t), e(t), t)R^{-1} \\ & \times \bar{G}^T(\tilde{x}(t), e(t), t)(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial [\tilde{x}^T(t) \;\; e^T(t)]^T}) \\ & + \frac{1}{4\rho^2}(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial [\tilde{x}^T(t) \;\; e^T(t)]^T})^T \begin{bmatrix} D(x(t)) & -L(\hat{x}(t)) \\ D_e(e(t), t) & 0 \end{bmatrix} \\ & \times \begin{bmatrix} D(x(t)) & -L(\hat{x}(t)) \\ D_e(e(t), t) & 0 \end{bmatrix}^T (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial [\tilde{x}^T(t) \;\; e^T(t)]^T}))dt\} \end{aligned} \tag{A5}$$

By the fact

$$\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial [\tilde{x}^T(t) \;\; e^T(t)]^T} = \begin{bmatrix} \frac{\partial V(\tilde{x}(t), e(t), t)}{\partial \tilde{x}(t)} \\ \frac{\partial V(\tilde{x}(t), e(t), t)}{\partial e(t)} \end{bmatrix} \tag{A6}$$

we get

$$\begin{aligned} & (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial [\tilde{x}^T(t) \;\; e^T(t)]^T})^T \begin{bmatrix} \tilde{F}(\tilde{x}(t)) - L(\hat{x}(t))\tilde{C}(\tilde{x}(t)) \\ F_e(e(t), t) \end{bmatrix} \\ & = (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial \tilde{x}(t)})^T \left(\tilde{F}(\tilde{x}(t)) - L(\hat{x}(t))\tilde{C}(\tilde{x}(t))\right) \\ & \quad + (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial e(t)})^T F_e(e(t), t) \end{aligned} \tag{A7}$$

and

$$\begin{aligned} & (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial [\tilde{x}^T(t) \;\; e^T(t)]^T})^T \begin{bmatrix} D(x(t)) & -L(\hat{x}(t)) \\ D_e(e(t), t) & 0 \end{bmatrix} \\ & \quad \times \begin{bmatrix} D(x(t)) & -L(\hat{x}(t)) \\ D_e(e(t), t) & 0 \end{bmatrix}^T (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial [\tilde{x}^T(t) \;\; e^T(t)]^T}) \\ & = \begin{bmatrix} (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial \tilde{x}(t)})^T & (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial e(t)})^T \end{bmatrix} \\ & \quad \times \begin{bmatrix} D(x(t))D^T(x(t)) & D(x(t))D_e^T(e(t), t) \\ +L(\hat{x}(t))L^T(\hat{x}(t)) & \\ D_e(e(t), t)D^T(x(t)) & D_e(e(t), t)D_e^T(e(t), t) \end{bmatrix} \\ & \quad \times \begin{bmatrix} (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial \tilde{x}(t)}) \\ (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial e(t)}) \end{bmatrix} = (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial \tilde{x}(t)})^T \\ & \quad \times \left(D(x(t))D^T(x(t)) + L(\hat{x}(t))L^T(\hat{x}(t))\right) \frac{\partial V(\tilde{x}(t), e(t), t)}{\partial \tilde{x}(t)} \\ & \quad + (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial \tilde{x}(t)})^T D(x(t), t)D_e^T(e(t), t) \\ & \quad \times (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial e(t)}) \\ & \quad + (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial e(t)})^T D_e(e(t), t)D^T(x(t)) \\ & \quad \times (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial \tilde{x}(t)}) \\ & \quad + (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial e(t)})^T D_e(e(t), t)D_e^T(e(t), t) \\ & \quad \times (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial e(t)}) \end{aligned} \tag{A8}$$

Substituting (A6), (A7) and (A8) into (A5), we get

$$\begin{aligned} J = & \min_{L(\hat{x}(t))} E\{-V(\tilde{x}(t_f), e(t_f), t_f) + V(\tilde{x}(0), e(0), 0) \\ & + \int_0^{t_f} (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial t} + \bar{x}^T(t)\bar{Q}\bar{x}(t) \\ & + (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial \tilde{x}(t)})^T \\ & \times \left(\tilde{F}(\tilde{x}(t)) - L(\hat{x}(t))\tilde{C}(\tilde{x}(t))\right) + (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial e(t)})^T \\ & \times F_e(e(t), t) - \frac{1}{4}(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial [\tilde{x}^T(t) \;\; e^T(t)]^T})^T \bar{G}(\tilde{x}(t), e(t), t)R^{-1} \\ & \times \bar{G}^T(\tilde{x}(t), e(t), t)(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial [\tilde{x}^T(t) \;\; e^T(t)]^T}) + \frac{1}{4\rho^2} \\ & \times (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial \tilde{x}(t)})^T (D(x(t))D^T(x(t)) \\ & + L(\hat{x}(t))L^T(\hat{x}(t)))(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial \tilde{x}(t)}) + \frac{1}{4\rho^2} \\ & \times (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial \tilde{x}(t)})^T D(x(t))D_e^T(e(t), t) \\ & \times (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial e(t)}) \\ & + \frac{1}{4\rho^2}(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial e(t)})^T D_e(e(t), t)D^T(x(t)) \\ & \times (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial \tilde{x}(t)}) + \frac{1}{4\rho^2}(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial e(t)})^T \\ & \times D_e(e(t), t) \\ & \times D_e^T(e(t), t)(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial e(t)}))dt\} \\ = & \min_{L(\hat{x}(t))} E\{-V(\tilde{x}(t_f), e(t_f), t_f) + V(\tilde{x}(0), e(0), 0) \\ & + \int_0^{t_f} (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial t} + \bar{x}^T(t)\bar{Q}\bar{x}(t) \\ & - \frac{1}{4}(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial [\tilde{x}^T(t) \;\; e^T(t)]^T})^T \bar{G}(\tilde{x}(t), e(t), t)R^{-1} \\ & \times \bar{G}(\tilde{x}(t), e(t), t)^T(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial [\tilde{x}^T(t) \;\; e^T(t)]^T}) \\ & + \left(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial \tilde{x}(t)}\right)^T \\ & \times \tilde{F}(\tilde{x}(t)) + (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial e(t)})^T F_e(e(t), t) + \frac{1}{4\rho^2} \\ & \times (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial \tilde{x}(t)})^T D(x(t))D^T(x(t))(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial \tilde{x}(t)}) \\ & + \frac{1}{4\rho^2}(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial \tilde{x}(t)})^T D(x(t))D_e^T(e(t), t) \\ & \times (\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial e(t)}) + \frac{1}{4\rho^2}(\frac{\partial V(\tilde{x}(t), e(t), t)}{\partial e(t)})^T \end{aligned}$$

$$\times D_e(e(t),t)$$
$$\times D^T(x(t))(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial \tilde{x}(t)})+\frac{1}{4\rho^2}(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial e(t)})^T$$
$$\times D_e(e(t),t)D_e^T(e(t),t)(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial e(t)})$$
$$+\frac{1}{4\rho^2}(L^T(\hat{x}(t))(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial \tilde{x}(t)})-\frac{1}{2}\tilde{C}(\tilde{x}(t)))^T$$
$$\times(L^T(\hat{x}(t))(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial \tilde{x}(t)})-\frac{1}{2}\tilde{C}(\tilde{x}(t)))$$
$$-\frac{1}{16\rho^2}\tilde{C}^T(\tilde{x}(t))\tilde{C}(\tilde{x}(t)))dt\} \quad \text{(A9)}$$

From (A9), we obtain the optimal $L^*(\hat{x}(t))$ as (15), i.e.,

$$\begin{aligned}J &= E\{-V(\tilde{x}(t_f),e(t_f),t_f)+V(\tilde{x}(0),e(0),0)\\ &+\int_0^{t_f}(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial t}+\bar{x}^T(t)\bar{Q}\bar{x}(t)\\ &-\frac{1}{4}(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)]^T})^T\bar{G}(\tilde{x}(t),e(t),t)R^{-1}\\ &\times\bar{G}(\tilde{x}(t),e(t),t)^T(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)]^T})\\ &+(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)]^T})^T\bar{F}(\tilde{x}(t),e(t),t)\\ &+\frac{1}{4\rho^2}(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)]^T})^T\bar{D}(\tilde{x}(t),e(t),t)\\ &\times(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)]^T})-\frac{1}{16\rho^2}\tilde{C}^T(\tilde{x}(t))\tilde{C}(\tilde{x}(t)))dt\}\end{aligned} \quad \text{(A10)}$$

(ii) Step 2:
By the HJIE of (16) and (A10), we get

$$J = E\left\{-V(\tilde{x}(t_f),e(t_f),t_f)+V(\tilde{x}(0),e(0),0)\right\} \quad \text{(A11)}$$

By the fact $V(\tilde{x}(t_f),e(t_f),t_f)\geq 0$, then

$$J \leq E\{V(\tilde{x}(0),e(0),0)\} \quad \text{(A12)}$$

(b) If $v(t)\in\mathcal{L}_2[0,\infty)$ and $n(t)\in\mathcal{L}_2[0,\infty)$, i.e., $\bar{v}(t)\in\mathcal{L}_2[0,\infty)$, from the definition of J in (12) with $t_f=\infty$, then (A12) becomes

$$\begin{aligned}J &= \min_{\substack{K(\hat{x}(t),e(t))\\ L(\hat{x}(t))}}\max_{\bar{v}(t)\in\mathcal{L}_2[0,\infty)}E\{\int_0^\infty(\bar{x}^T(t)\bar{Q}\bar{x}(t)\\ &\quad +u^T(t)Ru(t)-\rho\bar{v}^T(t)\bar{v}(t))dt\}\\ &= E\{\int_0^\infty(\bar{x}^T(t)\bar{Q}\bar{x}(t)+u^T(t)Ru(t)\\ &\quad -\rho\bar{v}^T(t)\bar{v}(t))dt\}\\ &\leq E\{V(\tilde{x}(0),e(0),0)\}\end{aligned} \quad \text{(A13)}$$

i.e.,

$$\begin{aligned}&E\{\int_0^\infty(\bar{x}^T(t)\bar{Q}\bar{x}(t)+u^T(t)Ru(t))dt\}\\ &\quad\leq E\{V(\tilde{x}(0),e(0),0)\}+\rho E\int_0^\infty\bar{v}^T(t)\bar{v}(t)dt\end{aligned} \quad \text{(A14)}$$

Since $\bar{v}^*(t)\in\mathcal{L}_2[0,\infty)$ and $V(\tilde{x}(0),e(0),0)$ are finite due to finite initial condition $\tilde{x}(0),e(0)$, the right hand side of (A14) is finite. Therefore, $\bar{x}(t)\to 0$ (or $\tilde{x}(t)\to 0$ and $e(t)\to 0$) and $u^*(t)\to 0$ as $t_f\to\infty$, i.e., the proposed H_∞ observer-based output feedback reference tracking control will achieve the mean-square asymptotical estimation and tracking if external disturbance $v(t)\in\mathcal{L}_2[0,\infty)$ and measurement noise $n(t)\in\mathcal{L}_2[0,\infty)$. Further, if reference signal $r(t)\in\mathcal{L}_2[0,\infty)$, i.e., $r(t)\to 0$ as $t_f\to\infty$, then $e(t)=x(t)-r(t)\to 0$ as $t_f\to\infty$ implies $x(t)\to 0$ as $t_f\to\infty$. Therefore, the whole closed loop system is mean-square asymptotically stable. The proof is completed.

APPENDIX B: PROOF OF THEOREM 2

(a) In (24), we suppose

$$(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T})_\varepsilon = (\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T})+\bar{h}(\tilde{x}(t),e(t),t) \quad \text{(B1)}$$

where $\bar{h}(\tilde{x}(t),e(t),t)$ is the error function between the approximate $(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T})_\varepsilon$ of $HJIE_\varepsilon=\varepsilon(\theta_s(t))$ in (24) and the real solution $\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T}$ of $HJIE=0$ in (16).

By substracting (24) from (16), $\varepsilon(\theta_s(t))$ could be rewritten as the follows:

$$\begin{aligned}&\varepsilon(\theta_s(t))\\ &= HJIE_\varepsilon - HJIE\\ &= \left((\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T})_\varepsilon^T-(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T})^T\right)\\ &\quad\times\begin{bmatrix}\bar{F}(\tilde{x}(t),e(t),t)\\ 1\end{bmatrix}-\frac{1}{4}(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T})_\varepsilon^T\\ &\quad\times\begin{bmatrix}\bar{G}(\tilde{x}(t),e(t),t)\\ 0\end{bmatrix}R^{-1}\begin{bmatrix}\bar{G}(\tilde{x}(t),e(t),t)\\ 0\end{bmatrix}^T\\ &\quad\times(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T})_\varepsilon+\frac{1}{4}(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T})^T\\ &\quad\times\begin{bmatrix}\bar{G}(\tilde{x}(t),e(t),t)\\ 0\end{bmatrix}R^{-1}\begin{bmatrix}\bar{G}(\tilde{x}(t),e(t),t)\\ 0\end{bmatrix}^T\\ &\quad\times(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T})+\frac{1}{4\rho^2}(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T})_\varepsilon^T\\ &\quad\times\begin{bmatrix}\bar{D}(\tilde{x}(t),e(t),t)&0\\ 0&0\end{bmatrix}(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T})_\varepsilon\\ &\quad-\frac{1}{4\rho^2}(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T})^T\begin{bmatrix}\bar{D}(\tilde{x}(t),e(t),t)&0\\ 0&0\end{bmatrix}\\ &\quad\times(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T})\end{aligned} \quad \text{(B2)}$$

By (B1), we get

$$\begin{aligned}&\varepsilon(\theta_s(t))\\ &= \bar{h}^T(\tilde{x}(t),e(t),t)\begin{bmatrix}\bar{F}(\tilde{x}(t),e(t),t)\\ 1\end{bmatrix}\\ &\quad-\frac{1}{4}\bar{h}^T(\tilde{x}(t),e(t),t)\begin{bmatrix}\bar{G}(\tilde{x}(t),e(t),t)\\ 0\end{bmatrix}R^{-1}\end{aligned}$$

$$\times\begin{bmatrix}\bar{G}(\tilde{x}(t),e(t),t)\\0\end{bmatrix}^T(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T})$$
$$-\frac{1}{4}(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T})^T\begin{bmatrix}\bar{G}(\tilde{x}(t),e(t),t)\\0\end{bmatrix}R^{-1}$$
$$\times\begin{bmatrix}\bar{G}(\tilde{x}(t),e(t),t)\\0\end{bmatrix}^T\bar{h}(\tilde{x}(t),e(t),t)$$
$$-\frac{1}{4}\bar{h}^T(\tilde{x}(t),e(t),t)\begin{bmatrix}\bar{G}(\tilde{x}(t),e(t),t)\\0\end{bmatrix}R^{-1}$$
$$\times\begin{bmatrix}\bar{G}(\tilde{x}(t),e(t),t)\\0\end{bmatrix}^T\bar{h}(\tilde{x}(t),e(t),t)$$
$$+\frac{1}{4\rho^2}\bar{h}^T(\tilde{x}(t),e(t),t)\begin{bmatrix}\bar{D}(\tilde{x}(t),e(t),t)&0\\0&0\end{bmatrix}$$
$$\times(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T})+\frac{1}{4\rho^2}(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T})^T$$
$$\times\begin{bmatrix}\bar{D}(\tilde{x}(t),e(t),t)&0\\0&0\end{bmatrix}\bar{h}(\tilde{x}(t),e(t),t)$$
$$+\frac{1}{4\rho^2}\bar{h}^T(\tilde{x}(t),e(t),t)\begin{bmatrix}\bar{D}(\tilde{x}(t),e(t),t)&0\\0&0\end{bmatrix}$$
$$\times\bar{h}(\tilde{x}(t),e(t),t)\tag{B3}$$

By the following symmetric properties

$$\bar{h}^T(\tilde{x}(t),e(t),t)\begin{bmatrix}\bar{G}(\tilde{x}(t),e(t),t)\\0\end{bmatrix}R^{-1}$$
$$\times\begin{bmatrix}\bar{G}(\tilde{x}(t),e(t),t)\\0\end{bmatrix}^T(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T})$$
$$=(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T})^T\begin{bmatrix}\bar{G}(\tilde{x}(t),e(t),t)\\0\end{bmatrix}R^{-1}$$
$$\times\begin{bmatrix}\bar{G}(\tilde{x}(t),e(t),t)\\0\end{bmatrix}^T\bar{h}(\tilde{x}(t),e(t),t)\tag{B4}$$

$$\bar{h}^T(\tilde{x}(t),e(t),t)\begin{bmatrix}\bar{D}(\tilde{x}(t),e(t),t)&0\\0&0\end{bmatrix}$$
$$\times(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T})=(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T})^T$$
$$\times\begin{bmatrix}\bar{D}(\tilde{x}(t),e(t),t)&0\\0&0\end{bmatrix}\bar{h}(\tilde{x}(t),e(t),t)\tag{B5}$$

By (B4) and (B5), then we get

$$\varepsilon(\theta_s(t))=\bar{h}^T(\tilde{x}(t),e(t),t)\{\begin{bmatrix}\bar{F}(\tilde{x}(t),e(t),t)\\1\end{bmatrix}$$
$$-\frac{1}{2}\begin{bmatrix}\bar{G}(\tilde{x}(t),e(t),t)\\0\end{bmatrix}R^{-1}$$
$$\times\begin{bmatrix}\bar{G}(\tilde{x}(t),e(t),t)\\0\end{bmatrix}^T(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T})$$
$$-\frac{1}{4}\begin{bmatrix}\bar{G}(\tilde{x}(t),e(t),t)\\0\end{bmatrix}R^{-1}$$
$$\times\begin{bmatrix}\bar{G}(\tilde{x}(t),e(t),t)\\0\end{bmatrix}^T\bar{h}(\tilde{x}(t),e(t),t)$$
$$+\frac{1}{2\rho^2}\begin{bmatrix}\bar{D}(\tilde{x}(t),e(t),t)&0\\0&0\end{bmatrix}(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T})$$
$$+\frac{1}{4\rho^2}\begin{bmatrix}\bar{D}(\tilde{x}(t),e(t),t)&0\\0&0\end{bmatrix}\bar{h}(\tilde{x}(t),e(t),t)\}\tag{B6}$$

If $\varepsilon(\theta_s(t))\to 0$ in (24), then (B6) becomes

$$\bar{h}^T(\tilde{x}(t),e(t),t)\{\begin{bmatrix}\bar{F}(\tilde{x}(t),e(t),t)\\1\end{bmatrix}$$
$$-\frac{1}{2}\begin{bmatrix}\bar{G}(\tilde{x}(t),e(t),t)\\0\end{bmatrix}R^{-1}\begin{bmatrix}\bar{G}(\tilde{x}(t),e(t),t)\\0\end{bmatrix}^T$$
$$\times(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T})-\frac{1}{4}\begin{bmatrix}\bar{G}(\tilde{x}(t),e(t),t)\\0\end{bmatrix}R^{-1}$$
$$\times\begin{bmatrix}\bar{G}(\tilde{x}(t),e(t),t)\\0\end{bmatrix}^T\bar{h}(\tilde{x}(t),e(t),t)$$
$$+\frac{1}{2\rho^2}\begin{bmatrix}\bar{D}(\tilde{x}(t),e(t),t)&0\\0&0\end{bmatrix}(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T})$$
$$+\frac{1}{4\rho^2}\begin{bmatrix}\bar{D}(\tilde{x}(t),e(t),t)&0\\0&0\end{bmatrix}\bar{h}(\tilde{x}(t),e(t),t)\}$$
$$\to 0\tag{B7}$$

Obviously, the term

$$\begin{bmatrix}\bar{F}(\tilde{x}(t),e(t),t)\\1\end{bmatrix}-\frac{1}{2}\begin{bmatrix}\bar{G}(\tilde{x}(t),e(t),t)\\0\end{bmatrix}R^{-1}$$
$$\times\begin{bmatrix}\bar{G}(\tilde{x}(t),e(t),t)\\0\end{bmatrix}^T(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T})$$
$$-\frac{1}{4}\begin{bmatrix}\bar{G}(\tilde{x}(t),e(t),t)\\0\end{bmatrix}R^{-1}\begin{bmatrix}\bar{G}(\tilde{x}(t),e(t),t)\\0\end{bmatrix}^T$$
$$\times\bar{h}(\tilde{x}(t),e(t),t)+\frac{1}{2\rho^2}\begin{bmatrix}\bar{D}(\tilde{x}(t),e(t),t)&0\\0&0\end{bmatrix}$$
$$\times(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T})+\frac{1}{4\rho^2}\begin{bmatrix}\bar{D}(\tilde{x}(t),e(t),t)&0\\0&0\end{bmatrix}$$
$$\times\bar{h}(\tilde{x}(t),e(t),t)$$

in $\{\cdot\}$ of (B7) is different from $HJIE=0$ in (16) and therefore is not equal to 0 for all $\tilde{x}(t),e(t)$. From (B7) we can conclude $\bar{h}(\tilde{x}(t),e(t),t)\to 0$. From (B1), it implies $(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T})_\varepsilon\to\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T}$, i.e., $HJIE_\varepsilon=\varepsilon(\theta_s(t))\to HJIE=0$ as $\varepsilon(\theta_s(t))\to 0$. According to Theorem 1, the HJIE-embedded DNN-based H_∞ observer-based output feedback reference tracking scheme in Fig. 1 can approach to the H_∞ observer-based output feedback reference tracking control in (14) and (15) in Theorem 1.

(b) Since we have shown $(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T})_\varepsilon\to\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)\ \ t]^T}$ and $HJIE_\varepsilon\to HJIE$ as $\varepsilon(\theta_s(t))\to 0$, the HJIE-embedded DNN-based robust H_∞ observer-based output feedback reference tracking control in Fig. 1 will approach to the theoretical H_∞ observer output feedback reference tracking control design (13)-(15), i.e., $\bar{v}(t)_\varepsilon=\frac{1}{2\rho^2}\begin{bmatrix}D(x(t))&-L^*(\hat{x}(t))\\D_e(e(t),t)&0\end{bmatrix}^T(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)]^T})_\varepsilon\to\bar{v}^*(t)$ in (13), $u(t)_\varepsilon=-\frac{1}{2}R^{-1}\begin{bmatrix}\tilde{G}(\tilde{x}(t))\\G_e(e(t),t)\end{bmatrix}^T\times(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial[\tilde{x}^T(t)\ \ e^T(t)]^T})_\varepsilon\to u^*(t)$ in (14) and $L(\hat{x}(t))_\varepsilon=\frac{1}{2}\times\frac{(\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial\tilde{x}(t)})}{\left\|\frac{\partial V(\tilde{x}(t),e(t),t)}{\partial\tilde{x}(t)}\right\|^2}$

$\tilde{C}^T(\tilde{x}(t)) \rightarrow L^*(\hat{x}(t))$ in (15). Then the remaining proof procedure of the asymptotical estimation, tracking and stability of closed loop system under $v(t) \in \mathcal{L}_2[0,\infty)$ and $n(t) \in \mathcal{L}_2[0,\infty)$ is similar to part (b) of Theorem 1. The proof is finished.

REFERENCES

[1] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, U.K.: MIT Press, 2016.

[2] G. E. Dahl, D. Yu, L. Deng, and A. Acero, "Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition," *IEEE Trans. Audio Speech Language Process.*, vol. 20, no. 1, pp. 30–42, Jan. 2012.

[3] A. L. Maas, P. Qi, Z. Xie, A. Y. Hannun, C. T. Lengerich, D. Jurafsky, and A. Y. Ng, "Building DNN acoustic models for large vocabulary speech recognition," *Comput. Speech Lang.*, vol. 41, pp. 195–213, Jan. 2017.

[4] D. Ciresan, U. Meier, and J. Schmidhuber, "Multi-column deep neural networks for image classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 3642–3649.

[5] H.-C. Shin, H.-C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, and R. M. Summers, "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1285–1298, May 2016.

[6] S. P. Singh, A. Kumar, H. Darbari, L. Singh, A. Rastogi, and S. Jain, "Machine translation using deep learning: An overview," in *Proc. Int. Conf. Comput. Commun. Electron.*, Jaipur, Rajasthan, 2017, pp. 162–167.

[7] C. Deng, S. Liao, Y. Xie, K. K. Parhi, X. Qian, and B. Yuan, "PermDNN: Efficient compressed DNN architecture with permuted diagonal matrices," in *Proc. 51st Annu. IEEE/ACM Int. Symp. Microarchitecture (MICRO)*, Fukuoka, Japan, Oct. 2018, pp. 189–202.

[8] A. E. Eshratifar, A. Esmaili, and M. Pedram, "BottleNet: A deep learning architecture for intelligent mobile cloud computing services," in *Proc. IEEE/ACM Int. Symp. Low Power Electron. Design (ISLPED)*, Lausanne, Switzerland, Jul. 2019, pp. 1–6.

[9] J. Yin, S. Su, J. Xun, T. Tang, and R. Liu, "Data-driven approaches for modeling train control models: Comparison and case studies," *ISA Trans.*, vol. 98, pp. 349–363, Mar. 2020.

[10] N. Berman and U. Shaked, "H_∞-like control for nonlinear stochastic control," *Syst. Control Lett.*, vol. 55, pp. 247–257, Mar. 2006.

[11] T. Basar and P. Bernhard, *H_∞ Optimal Control and Related Minmax Design Problems: A Dynamic Game Approach*. Boston, MA, USA: Birkhäuser, 1995.

[12] J. W. Helton and M. R. James, *Extending H_∞ Control to Nonlinear Systems*. Philadelphia, PA, USA: SIAM, 1999.

[13] A. Isidori and A. Astolfi, "Disturbance attenuation and H_∞ control via measurement feedback in nonlinear systems," *IEEE Trans. Autom. Control*, vol. 37, no. 9, pp. 1283–1293, Sep. 1992.

[14] J. Yong and X. Y. Zhou, *Stochastic Control: Hamiltonian Systems and HJB Equation*. New York, NY, USA: Springer, 1999.

[15] W. Zhang, L. Xie, and B. S. Chen, *Stochasitc H_2/H_∞ Control: A Nash Game Approach*. Boca Raton, FL, USA: CRC Press, 2017.

[16] C.-S. Tseng, B.-S. Chen, and H.-J. Uang, "Fuzzy tracking control design for nonlinear dynamic systems via T–S fuzzy model," *IEEE Trans. Fuzzy Syst.*, vol. 9, no. 3, pp. 381–392, Jun. 2001.

[17] B. S. Chen, C. P. Wang, and M. Y. Lee, "Stochastic robust team tracking control of multi-UAV networked system under Wiener and Poisson random fluctuations," *IEEE Trans. Cybern.*, vol. 51, no. 2, pp. 5786–5799, Dec. 2021.

[18] B. S. Chen, W. H. Chen, and H. L. Wu, "Robust H_2/H_∞ global linearization filter design for nonlinear stochastic systems," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 56, no. 7, pp. 1441–1454, Jul. 2009.

[19] K. A. Hoo and J. C. Kantor, "Global linearization and control of a mixed-culture bioreactor with competition and external inhibition," *Math. Biosci.*, vol. 82, no. 1, pp. 43–62, Nov. 1986.

[20] S. Boyd, L. Ei Ghaoui, E. Feron, and V. Balakrishan, *Linear Matrix Inequalities in System and Control Theory*. Philadelphia, PA, USA: SIAM, 1994.

[21] R. A. Nichols, R. T. Reichert, and W. J. Rugh, "Gain scheduling for H_∞ controllers: A flight control example," *IEEE Trans. Control Syst. Technol.*, vol. 1, no. 2, pp. 69–79, Jun. 1993.

[22] B.-S. Chen and C.-F. Wu, "Robust scheduling filter design for a class of nonlinear stochastic Poisson signal systems," *IEEE Trans. Signal Process.*, vol. 63, no. 23, pp. 6245–6257, Dec. 2015.

[23] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. ICLR*, San Diego, CA, USA, Dec. 2014, pp. 1–15.

[24] Y. N. Dauphin, H. D. Vries, J. Chung, and Y. Bengio, "RMSProp and equilibrated adaptive learning rates for non-convex optimization," Feb. 2015, *arXiv:1502.04390*.

[25] B. Zhao, B. Xian, Y. Zhang, and X. Zhang, "Nonlinear robust adaptive tracking control of a quadrotor UAV via immersion and invariance methodology," *IEEE Trans. Ind. Electron.*, vol. 62, no. 5, pp. 2891–2902, May 2015.

[26] M. Y. Lee, B. S. Chen, C. Y. Tsai, and C. L. Hwang, "Stochastic H_∞ robust decentralized tracking control of large-scale team formation UAV network system with time-varying delay and packet dropout under interconnected couplings and Wiener fluctuations," *IEEE Access*, vol. 9, pp. 41976–41997, 2021.

[27] B. S. Chen, M. Y. Lee, and T. H. Lin, "DNN-based H_∞ control scheme of nonlinear time-varying dynamic systems with external disturbance and its application to UAV tracking design," *IEEE Access*, vol. 9, pp. 69635–69653, 2021.

[28] Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Netw.*, vol. 2, no. 5, pp. 359–366, Jun. 1989.

[29] J. Li, Y. Fang, Y. Ge, and Y. Wu, "DNN-based implementation of data-driven iterative learning control for unknown system dynamics," in *Proc. IEEE 9th Data Driven Control Learn. Syst. Conf. (DDCLS)*, Nov. 2020, pp. 1037–1042.

[30] C. Wang, J. Wang, Y. Shen, and X. Zhang, "Autonomous navigation of UAVs in large-scale complex environments: A deep reinforcement learning approach," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 2124–2136, Mar. 2019.

[31] Z. Ning, K. Zhang, X. Wang, M. S. Obaidat, L. Guo, X. Hu, B. Hu, Y. Guo, B. Sadoun, and R. Y. K. Kwok, "Joint computing and caching in 5G-envisioned Internet of Vehicles: A deep reinforcement learning-based traffic control system," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 8, pp. 5201–5212, Aug. 2021.

BOR-SEN CHEN (Life Fellow, IEEE) received the B.S. degree in electrical engineering from the Tatung Institute of Technology, Taipei, Taiwan, in 1970, the M.S. degree in geophysics from the National Central University, Chungli, Taiwan, in 1973, and the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, CA, USA, in 1982. He has been a Lecturer, an Associate Professor, and a Professor at the Tatung Institute of Technology, from 1973 to 1987. He is currently the Tsing Hua Distinguished Chair Professor of electrical engineering and computer science at the National Tsing Hua University, Hsinchu, Taiwan. His current research interests include control engineering, signal processing, and systems biology. He has received the Distinguished Research Award from the National Science Council of Taiwan four times. He is a National Chair Professor of the Ministry of Education of Taiwan.

PO-HSUN WU received the B.S. degree in electrical engineering from the National Chung Cheng University, Chiayi, Taiwan, in 2018. He is currently pursuing the master's degree in electrical engineering with the National Tsing Hua University, Hsinchu, Taiwan. His current research interests include robust control, fuzzy control, and nonlinear stochastic systems.