2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Dynamic Neural Network-based Robust Observers for Second-order Uncertain Nonlinear Systems

H. Dinh, R. Kamalapurkar, S. Bhasin, and W. E. Dixon

***Abstract*— A dynamic neural network (DNN) based robust observer for second-order uncertain nonlinear systems is developed. The observer structure consists of a DNN to estimate the system dynamics on-line, a dynamic filter to estimate the unmeasurable state and a sliding mode feedback term to account for modeling errors and exogenous disturbances. The observed states are proven to asymptotically converge to the system states though Lyapunov-based stability analysis.**

I. INTRODUCTION

Full state feedback is not always available in many practical systems. In the absence of sensors, the requirement of full-state feedback for the controller is typically fulfilled by using *ad hoc* numerical differentiation techniques. The Euler difference algorithms are the simplest and the most common numerical methods, however, these approaches can aggravate the presence of noise and lead to unusable state estimates. The differentiation schemes proposed by Diop *et al.* in [1] and [2] are discrete and off-line methods. A more rigorous direction to estimate unmeasurable states in literature is nonlinear observer design. For instance, sliding observers were designed for general nonlinear systems by Slotine *et al.* in [3], for robot manipulators by Wit *et al.* in [4], and for mechanical systems subject to impacts by Mohamed *et al.* in [5]. However, all these observers require exact model knowledge to compensate for nonlinearities in the system. Model-based observers are also proposed in [6], [7] which require a high-gain to guarantee convergence of the estimation error. The observes introduced in [8] and [9] are both applied for Lagrangian dynamic systems to estimate the velocity, and asymptotic convergence to the true velocity is obtained. However, the symmetric positive-definiteness of the inertia matrix and the skew-symmetric property of the Coriolis matrix are required. Model knowledge is required in [8] and a partial differential equation needs to be solved to design observers. In [9], the system dynamics must be expressed in a non-minimal model and only mass and inertia parameters are unknown in the system.

Design of robust observers for uncertain nonlinear systems is considered in [10]–[12]. In [10], a second-order sliding mode observer for uncertain systems using super-twisting algorithm is proposed, where a nominal model of the system is assumed to be available and estimated errors are proven to converge in finite-time to a bounded set around the origin. In [11], the proposed observer can guarantee that the state estimates converge exponentially fast to the actual state, if there exists a vector function satisfying a complex set of matching conditions. In [12], the first asymptotic velocity observer for general second-order systems is proposed, where the estimation error is proven to asymptotically converge to zero. However, all nonlinear uncertainties in the system are damped out by a sliding mode term resulting in high frequency state estimates. A neural network (NN) approach that uses the universal approximation property is investigated for use in an adaptive observer design in [13]. However, estimation errors in this study are only guaranteed to be bounded due to function reconstruction inaccuracies.

The challenge to obtain asymptotic estimation stems from the fact that to robustly account for disturbances, feedback of the unmeasurable error and its estimate is required. Typically, feedback of the unmeasurable error is derived by taking the derivative of the measurable state and manipulating the resulting dynamics (e.g., this is the approach used in methods such as [12] and [13]). However, such an approach provides a linear feedback term of the unmeasurable state. Hence, a sliding mode term could not be simply added to the NN structure of the result in [13] to yield an asymptotic result, because it would require the signum of the unmeasurable state, and it does not seem clear how this nonlinear function of the unmeasurable state can be injected in the closed-loop error system using traditional methods. Likewise, it is not clear how to simply add a NN-based feedforward estimation of the nonlinearities in results such as [12] because of the need to inject nonlinear functions of the unmeasurable state. The novel approach used in this paper avoids this issue by using nonlinear (sliding mode) feedback of the measurable state, and then exploiting the recurrent nature of a dynamic neural network (DNN) structure to inject terms that cancel cross terms associated with the unmeasurable state. The approach is facilitated by using the filter structure of the controller in [12] and a novel stability analysis. The stability analysis is based on the idea of segregating the nonlinear uncertainties into terms which can be upper-bounded by constants and terms which can upper-bounded by states. The terms upper-bounded by states can be cancelled by the linear feedback of the measurable errors, while the terms upper-bounded by constants are partially rejected by the sign feedback (of the measurable state) and partially eliminated by the novel DNN-

[0]H. Dinh, R. Kamalapurkar, S. Bhasin, and W. E. Dixon are with the Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville FL 32611-6250, USA Email: {huyentdinh, rkamalapurkar, sbhasin, wdixon}@ufl.edu

[0]This research is supported in part by the NSF CAREER award number 0547448, NSF award number 0901491, the Department of Energy, grant number DE-FG04-86NE37967 as part of the DOE University Research Program in Robotics (URPR).

based weight update laws.

The contribution of this paper over previous results is that the observer is designed for uncertain nonlinear systems, and the on-line approximation of the unmeasurable uncertain nonlinearities via the DNN structure should heuristically improve the performance of methods that only use high-gain feedback. Asymptotic convergence of the estimated states to the real states is proven using a Lyapunov-based analysis. This observer can be used separately from the controller even if the relative degree between the control input and the output is arbitrary.

II. Robust Observer using Dynamic Neural Networks

Consider a second order control affine nonlinear system given by

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= f(x) + G(x)u + d, \\ y &= x_1, \end{aligned} \tag{1}$$

where $y(t) \in \mathbb{R}^n$ is the measurable output with a finite initial condition $y(0) = y_0$, $u(t) \in \mathbb{R}^m$ is the control input, $x(t) = [x_1(t)^T \ x_2(t)^T]^T \in \mathbb{R}^{2n}$ is the state of the system, $f(x) : \mathbb{R}^{2n} \to \mathbb{R}^n$, $G(x) : \mathbb{R}^{2n} \to \mathbb{R}^{n\times m}$ are unknown continuous functions, and $d(t) \in \mathbb{R}^n$ is an external disturbance. The following assumptions about the system in (1) will be utilized in the observer development.

Assumption 1: The state $x(t)$ is bounded, i.e, $x_1(t), x_2(t) \in \mathcal{L}_\infty$, and is partially measurable, i.e, $x_2(t)$ is unmeasurable.

Assumption 2: The unknown functions $f(x), G(x)$ and the control input $u(t)$ are C^1, and $u(t), \dot{u}(t) \in \mathcal{L}_\infty$.

Assumption 3: The disturbance $d(t)$ is differentiable, and $d(t), \dot{d}(t) \in \mathcal{L}_\infty$.

Assumption 4: Given a continuous function $F : \mathbb{S} \to \mathbb{R}^n$, where $\mathbb{S}$ is a simply connected compact set, there exist ideal weights $\theta = \theta^*$, such that the output of the NN, $\hat{F}(\cdot, \theta)$ approximates $F(\cdot)$ to an arbitrary accuracy [14].

Based on the universal approximation property of the multilayer NNs (MLNN) [15], [16], using Assumption 4, the unknown functions $f(x), G(x)$ in (1) can be replaced by multi-layer NNs (MLNN) as

$$\begin{aligned} f(x) &= W_f^T \sigma_f(V_{f_1}^T x_1 + V_{f_2}^T x_2) + \varepsilon_f(x), \\ g_i(x) &= W_{gi}^T \sigma_{gi}(V_{gi_1}^T x_1 + V_{gi_2}^T x_2) + \varepsilon_{gi}(x), \end{aligned}$$

where $W_f \in \mathbb{R}^{N_f+1\times n}$, $V_{f_1}, V_{f_2} \in \mathbb{R}^{n\times N_f}$ are unknown ideal weight matrices of the MLNN having N_f hidden layer neurons, $g_i(x)$ is the i^{th} column of the matrix $G(x)$, $W_{gi} \in \mathbb{R}^{N_{gi}+1\times n}$, $V_{gi_1}, V_{gi_2} \in \mathbb{R}^{n\times N_{gi}}$ are also unknown ideal weight matrices of the MLNN having N_{gi} hidden layer neurons, $i = 1...m$, $\sigma_f(t) \triangleq \sigma_f(V_{f_1}^T x_1(t) + V_{f_2}^T x_2(t)) \in \mathbb{R}^{N_f+1}$, $\sigma_{gi}(t) \triangleq \sigma_{gi}(V_{gi_1}^T x_1(t) + V_{gi_2}^T x_2(t)) \in \mathbb{R}^{N_{gi}+1}$ are the activation functions (sigmoid, hyperbolic tangent etc.), $\varepsilon_f(x), \varepsilon_{gi}(x) \in \mathbb{R}^n, i = 1...m$, are the function reconstruction errors. Hence, the system in (1) can be represented as

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= W_f^T \sigma_f + \varepsilon_f(x) + d \\ &\quad + \sum_{i=1}^m \left[W_{gi}^T \sigma_{gi} + \varepsilon_{gi}(x) \right] u_i, \end{aligned} \tag{2}$$

where $u_i(t) \in \mathbb{R}$ is the i^{th} element of the control input vector $u(t)$. The following assumptions will be used in the observer development and stability analysis.

Assumption 5: The ideal NN weights are bounded by known positive constants [17], i.e. $\|W_f\| \le \bar{W}_f$, $\|V_{f_1}\| \le \bar{V}_{f_1}$, $\|V_{f_2}\| \le \bar{V}_{f_2}$, $\|W_{gi}\| \le \bar{W}_{gi}$, $\|V_{gi_1}\| \le \bar{V}_{gi_1}$, and $\|V_{gi_2}\| \le \bar{V}_{gi_2}$, $i = 1...m$, where $\|\cdot\|$ denotes Frobenius norm for a matrix and Euclidean norm for a vector.

Assumption 6: The activation functions $\sigma_f(\cdot), \sigma_{gi}(\cdot)$ and their derivatives with respect to its arguments, $\sigma'_f(\cdot), \sigma'_{gi}(\cdot)$, $\sigma''_f(\cdot), \sigma''_{gi}(\cdot)$, $i = 1...m$, are bounded.

Assumption 7: The function reconstruction errors $\varepsilon_f(\cdot), \varepsilon_{gi}(\cdot)$, and its first derivatives with respect to their arguments are bounded, with $i = 1...m$ [17].

Remark 1: Assumptions 5-7 are standard assumptions in NN control literature (For details, see [17]). The idea weights are unknown and may not even unique, however, their existence is only required. The upper bounds of the ideal weights are assumed to be known to exploit in the projection algorithm to ensure that the DNN weight estimates are always bounded. Activation functions chosen as sigmoid, hyperbolic tangent functions satisfy Assumption 6.

The following multi-layer dynamic neural network (MLDNN) architecture is proposed to observe the system in (1)

$$\begin{aligned} \dot{\hat{x}}_1 &= \hat{x}_2, \\ \dot{\hat{x}}_2 &= \hat{W}_f^T \hat{\sigma}_f + \sum_{i=1}^m \hat{W}_{gi}^T \hat{\sigma}_{gi} u_i + v, \end{aligned} \tag{3}$$

where $\hat{x}(t)=[\hat{x}_1(t)^T \ \hat{x}_2(t)^T]^T \in \mathbb{R}^{2n}$ is the state of the DNN observer, $\hat{W}_f(t) \in \mathbb{R}^{N_f+1\times n}$, $\hat{V}_{f_1}(t), \hat{V}_{f_2}(t) \in \mathbb{R}^{n\times N_f}$, $\hat{W}_{gi}(t) \in \mathbb{R}^{N_{gi}+1\times n}$, $\hat{V}_{gi_1}(t), \hat{V}_{gi_2}(t) \in \mathbb{R}^{n\times N_{gi}}$, $i = 1...m$, are the weight estimates, $\hat{\sigma}_f(t) \triangleq \sigma_f(\hat{V}_{f_1}(t)^T \hat{x}_1(t) + \hat{V}_{f_2}(t)^T \hat{x}_2(t)) \in \mathbb{R}^{N_f+1}, \hat{\sigma}_{gi}(t) \triangleq \sigma_{gi}(\hat{V}_{gi_1}(t)^T \hat{x}_1(t) + \hat{V}_{gi_2}(t)^T \hat{x}_2(t)) \in \mathbb{R}^{N_{gi}+1}$, and $v(t) \in \mathbb{R}^n$ is a function to be determined to provide robustness to account for the function reconstruction errors and external disturbances.

The objective in this paper is to prove that the estimated state $\hat{x}(t)$ converges to the system state $x(t)$. To facilitate the subsequent analysis, the estimation error $\tilde{x}(t) \in \mathbb{R}^n$ is defined as

$$\tilde{x} \triangleq x_1 - \hat{x}_1. \tag{4}$$

To compensate for the lack of direct measurements of $x_2(t)$, a filtered estimation error is defined as

$$r \triangleq \dot{\tilde{x}} + \alpha \tilde{x} + \eta, \tag{5}$$

where $\alpha \in \mathbb{R}$ is a positive constant control gain, and $\eta(t) \in \mathbb{R}^n$ is an output of the dynamic filter [12]

$$\begin{aligned} \dot{p} &= -(k+2\alpha)p - \tilde{x}_f + ((k+\alpha)^2+1)\tilde{x}, \quad (6) \\ \dot{\tilde{x}}_f &= p - \alpha\tilde{x}_f - (k+\alpha)\tilde{x}, \quad (7) \\ p(0) &= (k+\alpha)\tilde{x}(0), \quad \tilde{x}_f(0) = 0, \\ \eta &= p - (k+\alpha)\tilde{x}, \quad (8) \end{aligned}$$

where $\tilde{x}_f(t) \in \mathbb{R}^n$ is another output of the filter, $p(t) \in \mathbb{R}^n$ is used as internal filter variable, and $k \in \mathbb{R}$ is a positive constant gain. The filtered estimation error $r(t)$ is not measurable, since the expression in (5) depend on $\dot{x}(t)$.

Remark 2: The second order dynamic filter to estimate the system velocity was first proposed for the output feedback control in [12]. The filter in (6)-(8) admits the estimation error $\tilde{x}(t)$ as its input and produces two signal outputs $\tilde{x}_f(t)$ and $\eta(t)$. The auxiliary signal $p(t)$ is utilized to only generate the signal $\eta(t)$ without involving the derivative of the estimation error $\dot{\tilde{x}}(t)$ which is unmeasurable. Hence, the filter can be physically implemented. A difficulty to obtain asymptotic estimation is that the filtered estimation error $r(t)$ is not available for feedback. The relation between two filter outputs is $\eta = \dot{\tilde{x}}_f + \alpha\tilde{x}_f$, and this relationship is utilized to generate the feedback of $r(t)$. Since taking time derivative of $r(t)$, the term $\ddot{x}_f(t)$ appears implicitly inside $\dot{\eta}(t)$, and consequently, the unmeasurable term $\dot{\tilde{x}}(t)$ which can be replaced by $r(t)$ is introduced.

Taking the derivative of (8) and using the definitions (5)-(8) yields

$$\dot{\eta} = -(k+\alpha)r - \alpha\eta + \tilde{x} - \tilde{x}_f. \quad (9)$$

The closed-loop dynamics of the derivative of the filtered estimation error in (5) is calculated by using (2)-(5), and (9) as

$$\begin{aligned} \dot{r} &= W_f^T\sigma_f - \hat{W}_f^T\hat{\sigma}_f + \sum_{i=1}^{m}[W_{gi}^T\sigma_{gi} - \hat{W}_{gi}^T\hat{\sigma}_{gi}]u_i \\ &\quad +\varepsilon_f + \sum_{i=1}^{m}\varepsilon_{gi}u_i + d - v + \alpha(r - \alpha\tilde{x} - \eta) \\ &\quad -(k+\alpha)r - \alpha\eta + \tilde{x} - \tilde{x}_f. \quad (10) \end{aligned}$$

The robustifing term $v(t)$ is designed based on the subsequent analysis as

$$\begin{aligned} v &= -[\gamma(k+\alpha)+2\alpha]\eta + (\gamma - \alpha^2)\tilde{x} \\ &\quad +\beta_1 sgn(\tilde{x} + \tilde{x}_f), \quad (11) \end{aligned}$$

where $\gamma, \beta_1 \in \mathbb{R}$ are positive constant control gains. Adding and subtracting $W_f^T\sigma_f(\hat{V}_{f_1}^T x_1 + \hat{V}_{f_2}x_2) + \hat{W}_f^T\sigma_f(\hat{V}_{f_1}^T x_1 + \hat{V}_{f_2}^T x_2) + \sum_{i=1}^{m}[W_{gi}^T\sigma_{gi}(\hat{V}_{gi_1}^T x_1 + \hat{V}_{gi_2}^T x_2) + \hat{W}_{gi}^T\sigma_{gi}(\hat{V}_{gi_1}^T x_1 + \hat{V}_{gi_2}^T x_2)]u_i$ and substituting $v(t)$ from (11), the expression in (10) can be rewritten as

$$\dot{r} = \tilde{N} + N - kr - \beta_1 sgn(\tilde{x} + \tilde{x}_f) + \gamma(k+\alpha)\eta - \gamma\tilde{x}, \quad (12)$$

where the auxiliary function $\tilde{N}(x_1, x_2, \hat{x}_1, \hat{x}_2, \tilde{x}_f, \hat{W}_f, \hat{V}_{f_1}, \hat{V}_{f_2}, \hat{W}_{gi}, \hat{V}_{gi_1}, \hat{V}_{gi_2}, t) \in \mathbb{R}^n$ is defined as

$$\begin{aligned} \tilde{N} &\triangleq \hat{W}_f^T[\sigma_f(\hat{V}_{f_1}^T x_1 + \hat{V}_{f_2}^T x_2) - \hat{\sigma}_f] + \tilde{x} - \tilde{x}_f \\ &\quad + \sum_{i=1}^{m}\hat{W}_{gi}^T[\sigma_{gi}(\hat{V}_{gi_1}^T x_1 + \hat{V}_{gi_2}^T x_2) - \hat{\sigma}_{gi}]u_i, \quad (13) \end{aligned}$$

and $N(x_1, x_2, \hat{W}_f, \hat{V}_{f_1}, \hat{V}_{f_2}, \hat{W}_{gi}, \hat{V}_{gi_1}, \hat{V}_{gi_2}, t) \in \mathbb{R}^n$ is segregated into two parts as

$$N \triangleq N_1 + N_2. \quad (14)$$

In (14), $N_1(x_1, x_2, \hat{W}_f, \hat{V}_{f_1}, \hat{V}_{f_2}, \hat{W}_{gi}, \hat{V}_{gi_1}, \hat{V}_{gi_2}, t)$, $N_2(x_1, x_2, \hat{W}_f, \hat{V}_{f_1}, \hat{V}_{f_2}, \hat{W}_{gi}, \hat{V}_{gi_1}, \hat{V}_{gi_2}, t) \in \mathbb{R}^n$ are defined as

$$\begin{aligned} N_1 &\triangleq \tilde{W}_f^T\sigma'_f[\tilde{V}_{f_1}^T x_1 + \tilde{V}_{f_2}^T x_2] + W_f^T O(\tilde{V}_{f_1}^T x_1 + \tilde{V}_{f_2}^T x_2)^2 \\ &\quad + \sum_{i=1}^{m}\tilde{W}_{gi}^T\sigma'_{gi}[\tilde{V}_{gi_1}^T x_1 + \tilde{V}_{gi_2}^T x_2]u_i \\ &\quad + \sum_{i=1}^{m}W_{gi}^T O(\tilde{V}_{gi_1}^T x_1 + \tilde{V}_{gi_2}^T x_2)^2 u_i \\ &\quad +\varepsilon_f + \sum_{i=1}^{m}\varepsilon_{gi}u_i + d, \end{aligned}$$

$$\begin{aligned} N_2 &\triangleq \tilde{W}_f^T\sigma_f(\hat{V}_{f_1}^T x_1 + \hat{V}_{f_2}^T x_2) + \hat{W}_f^T\sigma'_f[\tilde{V}_{f_1}^T x_1 + \tilde{V}_{f_2}^T x_2] \\ &\quad + \sum_{i=1}^{m}\tilde{W}_{gi}^T\sigma_{gi}(\hat{V}_{gi_1}^T x_1 + \hat{V}_{gi_2}^T x_2)u_i \\ &\quad + \sum_{i=1}^{m}\hat{W}_{gi}^T\sigma'_{gi}[\tilde{V}_{gi_1}^T x_1 + \tilde{V}_{gi_2}^T x_2]u_i, \quad (15) \end{aligned}$$

where $\tilde{W}_f(t) \triangleq W_f - \hat{W}_f(t) \in \mathbb{R}^{N_f+1\times n}$, $\tilde{V}_{f_1}(t) \triangleq V_{f_1} - \hat{V}_{f_1}(t) \in \mathbb{R}^{n\times N_f}$, $\tilde{V}_{f_2}(t) \triangleq V_{f_2} - \hat{V}_{f_2}(t) \in \mathbb{R}^{n\times N_f}$, $\tilde{W}_{gi}(t) \triangleq W_{gi} - \hat{W}_{gi}(t) \in \mathbb{R}^{N_{gi}+1\times n}$, $\tilde{V}_{gi_1}(t) \triangleq V_{gi_1} - \hat{V}_{gi_1}(t) \in \mathbb{R}^{n\times N_{gi}}$, $\tilde{V}_{gi_2}(t) \triangleq V_{gi_2} - \hat{V}_{gi_2}(t) \in \mathbb{R}^{n\times N_{gi}}$, $i = 1...m$, are the estimate mismatches for the ideal NN weights; $O(\tilde{V}_{f_1}^T x_1 + \tilde{V}_{f_2}^T x_2)^2(t) \in \mathbb{R}^{N_f+1}$, $O(\tilde{V}_{gi_1}^T x_1 + \tilde{V}_{gi_2}^T x_2)^2(t) \in \mathbb{R}^{N_{gi}+1}$ are the higher order terms in the Taylor series of the vector functions $\sigma_f(\cdot)$, $\sigma_{gi}(\cdot)$ in the neighborhood of $\hat{V}_{f_1}^T x_1 + \hat{V}_{f_2}^T x_2$ and $\hat{V}_{gi_1}^T x_1 + \hat{V}_{gi_2}^T x_2$, respectively, as

$$\begin{aligned} \sigma_f &= \sigma_f(\hat{V}_{f_1}^T x_1 + \hat{V}_{f_2}^T x_2) + \sigma'_f[\tilde{V}_{f_1}^T x_1 + \tilde{V}_{f_2}^T x_2] \\ &\quad +O(\tilde{V}_{f_1}^T x_1 + \tilde{V}_{f_2}^T x_2)^2, \quad (16) \\ \sigma_{gi} &= \sigma_{gi}(\hat{V}_{gi_1}^T x_1 + \hat{V}_{gi_2}^T x_2) + \sigma'_{gi}[\tilde{V}_{gi_1}^T x_1 + \tilde{V}_{gi_2}^T x_2] \\ &\quad +O(\tilde{V}_{gi_1}^T x_1 + \tilde{V}_{gi_2}^T x_2)^2, \end{aligned}$$

where the terms $\sigma'_f(t)$, $\sigma'_{gi}(t)$ are defined as $\sigma'_f(t) \triangleq \sigma'_f(\hat{V}_{f_1}(t)^T x_1(t) + \hat{V}_{f_2}(t)^T x_2(t)) = d\sigma_f(\theta)/d\theta|_{\theta=\hat{V}_{f_1}^T x_1 + \hat{V}_{f_2}^T x_2}$ and $\sigma'_{gi}(t) \triangleq \sigma'_{gi}(\hat{V}_{gi_1}(t)^T x_1(t) + \hat{V}_{gi_2}(t)^T x_2(t)) = d\sigma_{gi}(\theta)/d\theta|_{\theta=\hat{V}_{gi_1}^T x_1 + \hat{V}_{gi_2}^T x_2}$. To facilitate the subsequent analysis, an auxiliary function $\tilde{N}_2(\hat{x}_1, \hat{x}_2, \hat{W}_f, \hat{V}_{f_1}, \hat{V}_{f_2}, \hat{W}_{gi}, \hat{V}_{gi_1}, \hat{V}_{gi_2}, t) \in \mathbb{R}^n$ is defined by replacing terms $x_1(t), x_2(t)$ in $N_2(\cdot)$ by $\hat{x}_1(t), \hat{x}_2(t)$, respectively.

The weight update laws for the DNN in (3) are developed based on the subsequent stability analysis as

$$\begin{aligned}
\dot{\hat{W}}_f &= proj[\Gamma_{wf}\hat{\sigma}_f(\tilde{x}+\tilde{x}_f)^T], \\
\dot{\hat{V}}_{f_1} &= proj[\Gamma_{vf_1}\hat{x}_1(\tilde{x}+\tilde{x}_f)^T\hat{W}_f^T\hat{\sigma}'_f], \\
\dot{\hat{V}}_{f_2} &= proj[\Gamma_{vf_2}\hat{x}_2(\tilde{x}+\tilde{x}_f)^T\hat{W}_f^T\hat{\sigma}'_f], \\
\dot{\hat{W}}_{gi} &= proj[\Gamma_{wgi}\hat{\sigma}_{gi}u_i(\tilde{x}+\tilde{x}_f)^T], \qquad i=1...m \\
\dot{\hat{V}}_{gi_1} &= proj[\Gamma_{vgi_1}\hat{x}_1u_i(\tilde{x}+\tilde{x}_f)^T\hat{W}_{gi}^T\hat{\sigma}'_{gi}], \;\; i=1...m \\
\dot{\hat{V}}_{gi_2} &= proj[\Gamma_{vgi_2}\hat{x}_2u_i(\tilde{x}+\tilde{x}_f)^T\hat{W}_{gi}^T\hat{\sigma}'_{gi}], \;\; i=1...m
\end{aligned} \tag{17}$$

where $\Gamma_{wf} \in \mathbb{R}^{(N_f+1)\times(N_f+1)}$, $\Gamma_{wgi} \in \mathbb{R}^{(N_{gi}+1)\times(N_{gi}+1)}$, $\Gamma_{vf_1}, \Gamma_{vf_2}, \Gamma_{vgi_1}, \Gamma_{vgi_2} \in \mathbb{R}^{n\times n}$, are constant symmetric positive-definite adaptation gains, the terms $\hat{\sigma}'_f(t), \hat{\sigma}'_{gi}(t)$ are defined as $\hat{\sigma}'_f(t) \triangleq \sigma'_f(\hat{V}_{f_1}(t)^T\hat{x}_1(t) + \hat{V}_{f_2}(t)^T\hat{x}_2(t)) = d\sigma_f(\theta)/d\theta|_{\theta=\hat{V}_{f_1}^T\hat{x}_1+\hat{V}_{f_2}^T\hat{x}_2}$, $\hat{\sigma}'_{gi}(t) \triangleq \sigma'_{gi}(\hat{V}_{gi_1}(t)^T\hat{x}_1(t) + \hat{V}_{gi_2}(t)^T\hat{x}_2(t)) = d\sigma_{gi}(\theta)/d\theta|_{\theta=\hat{V}_{gi_1}^T\hat{x}_1+\hat{V}_{gi_2}^T\hat{x}_2}$, and $proj(\cdot)$ is a smooth projection operator [18], [19] used to guarantee that the weight estimates $\hat{W}_f(t), \hat{V}_{f_1}(t), \hat{V}_{f_2}(t), \hat{W}_{gi}(t), \hat{V}_{gi_1}(t)$, and $\hat{V}_{gi_2}(t)$ remain bounded.

Remark 3: In (3), the feedforward NN terms $\hat{W}_f(t)^T\hat{\sigma}_f(t), \hat{W}_{gi}(t)^T\hat{\sigma}_{gi}(t)$ are fed by the observer states $\hat{x}(t)$, hence this observer has a DNN structure. The DNN has a recurrent feedback loop, and is proven to be able to approximate dynamic systems with any arbitrarily small degree of accuracy [14], [20]. This property motivates for the DNN-based observer design. The DNN is tuned on-line to mimic the system dynamics by the weight update laws based on the state, weight estimates, and the filter output.

Using (4)-(8), Assumptions $1-2, 5-6$, the $proj(\cdot)$ algorithm in (17) and the Mean Value Theorem [21], the auxiliary function $\tilde{N}(\cdot)$ in (13) can be upper-bounded as

$$\left\|\tilde{N}\right\| \leq \zeta_1 \left\|z\right\|, \tag{18}$$

where $z(t) \in \mathbb{R}^{4n}$ is defined as

$$z(t) \triangleq [\tilde{x}^T \;\; \tilde{x}_f^T \;\; \eta^T \;\; r^T]^T. \tag{19}$$

Based on (4)-(8), Assumptions $1-3, 5-7$, the Taylor series expansion in (16) and the weight update laws in (17), the following bounds can be developed

$$\begin{aligned}
\|N_1\| &\leq \zeta_2, \;\; \|N_2\| \leq \zeta_3, \\
\left\|\dot{N}\right\| &\leq \zeta_4 + \rho(\|z\|)\,\|z\|, \\
\left\|\tilde{N}_2\right\| &\leq \zeta_5 \|z\|,
\end{aligned} \tag{20}$$

where $\zeta_i \in \mathbb{R}, i = 1...5$, are computable positive constants, $\rho(\cdot) \in \mathbb{R}$ is a positive, globally invertible, non-decreasing function, and $\tilde{N}_2(\tilde{x}, \dot{\tilde{x}}, \hat{W}_f, \hat{V}_{f_1}, \hat{V}_{f_2}, \hat{W}_{gi}, \hat{V}_{gi_1}, \hat{V}_{gi_2}, u) \triangleq N_2(\cdot) - \hat{N}_2(\cdot)$.

To facilitate the subsequent stability analysis, let $\mathcal{D} \subset \mathbb{R}^{4n+2}$ be a domain containing $y(t) = 0$, where $y(t) \in \mathbb{R}^{4n+2}$ is defined as

$$y(t) \triangleq [z^T(t) \;\; \sqrt{P(t)} \;\; \sqrt{Q(t)}]^T. \tag{21}$$

In (21), the auxiliary function $P(t) \in \mathbb{R}$ is the generated solution to the differential equation as

$$\begin{aligned}
\dot{P}(t) &\triangleq -L(t), && (22)\\
P(0) &\triangleq \beta_1 \sum_{j=1}^{n} \left|\tilde{x}_j(0) + \tilde{x}_{f_j}(0)\right| - (\tilde{x}(0)+\tilde{x}_f(0))^T N(0),
\end{aligned}$$

where the subscript $j = 1, 2, .., n$ denotes the j^{th} element of $\tilde{x}(0)$ or $\tilde{x}_f(0)$, and the auxiliary function $L(t) \in \mathbb{R}$ is defined as

$$\begin{aligned}
L(t) \;\triangleq\; & r^T(N_1 - \beta_1 sgn(\tilde{x}+\tilde{x}_f)) + (\dot{\tilde{x}}+\dot{\tilde{x}}_f)^T N_2 \\
& -\sqrt{2}\rho(\|z\|)\,\|z\|^2,
\end{aligned} \tag{23}$$

where $\beta_1 \in \mathbb{R}$ is a positive constant chosen according to the sufficient condition

$$\beta_1 > \max(\zeta_2+\zeta_3, \zeta_2+\frac{\zeta_4}{\alpha}), \tag{24}$$

where $\zeta_i, i = 2, 3, 4$ are introduced in (20). Provided the sufficient condition in (24) is satisfied, the following inequality can be obtained $P(t) \geq 0$ (see [22], [21]). The auxiliary function $Q(t) \in \mathbb{R}$ in (21) is defined as

$$\begin{aligned}
Q(t) \;\triangleq\; & \frac{\alpha}{2}tr(\tilde{W}_f^T\Gamma_{wf}^{-1}\tilde{W}_f) + \frac{\alpha}{2}tr(\tilde{V}_{f_1}^T\Gamma_{vf_1}^{-1}\tilde{V}_{f_1}) \\
& +\frac{\alpha}{2}tr(\tilde{V}_{f_2}^T\Gamma_{vf_2}^{-1}\tilde{V}_{f_2}) + \frac{\alpha}{2}\sum_{i=1}^{m} tr(\tilde{W}_{gi}^T\Gamma_{wg_i}^{-1}\tilde{W}_{gi}) \\
& +\frac{\alpha}{2}\sum_{i=1}^{m} tr(\tilde{V}_{gi_1}^T\Gamma_{vgi_1}^{-1}\tilde{V}_{gi_1}) \\
& +\frac{\alpha}{2}\sum_{i=1}^{m} tr(\tilde{V}_{gi_2}^T\Gamma_{vgi_2}^{-1}\tilde{V}_{gi_2}),
\end{aligned} \tag{25}$$

where $tr(\cdot)$ denotes the trace of a matrix. Since the gains $\Gamma_{wf}, \Gamma_{wgi}, \Gamma_{vf_1}, \Gamma_{vf_2}, \Gamma_{vgi_1}, \Gamma_{vgi_2}$ are symmetric, positive-definite matrices, $Q(t) \geq 0$.

III. Lyapunov Stability Analysis For DNN-based Observer

Theorem 1: The dynamic neural network-based observer proposed in (3) along with its weight update laws in (17) ensures asymptotic estimation in sense that

$$\|\tilde{x}(t)\| \to 0 \text{ and } \left\|\dot{\tilde{x}}(t)\right\| \to 0 \text{ as } t \to \infty$$

provided the control gain $k = k_1 + k_2$ introduced in (6)-(8) is selected sufficiently large based on the initial conditions of the states (see the subsequent proof), the gain condition in (24) is satisfied, and the following sufficient conditions are satisfied

$$\gamma > \alpha\zeta_5^2 + \frac{1}{2\alpha}, \;\; k_1 > \frac{1}{2}, \text{ and } \lambda > \frac{\zeta_1^2}{4\sqrt{2}k_2}, \tag{26}$$

where

$$\lambda \triangleq \frac{1}{\sqrt{2}}\left[\min(\alpha(\gamma-\alpha\zeta_5^2),k_1)-\frac{1}{2}\right], \tag{27}$$

and ζ_1, ζ_5 are introduced in (18), (20), respectively.

Proof: Consider the Lyapunov candidate function $V_L(y):\mathcal{D}\to\mathbb{R}$, which is a Lipschitz continuous regular positive definite function defined as

$$V_L \triangleq \frac{\gamma}{2}\tilde{x}^T\tilde{x}+\frac{\gamma}{2}\tilde{x}_f^T\tilde{x}_f+\frac{\gamma}{2}\eta^T\eta+\frac{1}{2}r^Tr+P+Q, \tag{28}$$

which satisfies the following inequalities:

$$U_1(y)\le V_L(y)\le U_2(y). \tag{29}$$

In (29), $U_1(y), U_2(y)\in\mathbb{R}$ are continuous positive definite functions defined as

$$U_1(y)\triangleq\varepsilon_1\|y\|^2,\ U_2(y)\triangleq\varepsilon_2\|y\|^2,$$

where $\varepsilon_1,\varepsilon_2\in\mathbb{R}$ are defined as

$$\varepsilon_1\triangleq\min(\frac{\gamma}{2},\frac{1}{2}),\ \varepsilon_2\triangleq\max(\frac{\gamma}{2},1).$$

Using (5), (12), (22), the differential equations of the closed-loop system are continuous except in the set $\{y|\tilde{x}=0\}$. Using Filippov's differential inclusion [23]–[26], the existence of solutions can be established for $\dot{y}=f(y)$, where $f(y)\in\mathbb{R}^{4n+2}$ denotes the right-hand side of the closed-loop error signals. Under Filippov's framework, a generalized Lyapunov stability theory can be used (see [26]–[29] for further details). The generalized time derivative of (28) exists almost everywhere (a.e.), and $\dot{V}(y)\in^{a.e.}\dot{\tilde{V}}(y)$ where

$$\dot{\tilde{V}}=\underset{\xi\in\partial V(y)}{\cap}\xi^TK\left[\dot{\tilde{x}}^T\ \dot{\tilde{x}}_f^T\ \dot{\eta}^T\ \dot{r}^T\ \frac{1}{2}P^{-\frac{1}{2}}\dot{P}\ \frac{1}{2}Q^{-\frac{1}{2}}\dot{Q}\right]^T,$$

where ∂V is the generalized gradient of $V(y)$ [27], and $K[\cdot]$ is defined as [28], [29]

$$K[f](y)\triangleq\underset{\delta>0}{\cap}\underset{\mu M=0}{\cap}\overline{co}f(B(y,\delta)-M),$$

where $\underset{\mu M=0}{\cap}$ denotes the intersection of all sets M of Lebesgue measure zero, $\overline{co}$ denotes convex closure, and $B(y,\delta)=\left\{w\in R^{4n+2}|\ \|y-w\|<\delta\right\}$. Since $V(y)$ is a Lipschitz continuous regular function,

$$\begin{aligned}\dot{\tilde{V}} &= \nabla V^TK\left[\dot{\tilde{x}}^T\ \dot{\tilde{x}}_f^T\ \dot{\eta}^T\ \dot{r}^T\ \frac{1}{2}P^{-\frac{1}{2}}\dot{P}\ \frac{1}{2}Q^{-\frac{1}{2}}\dot{Q}\right]^T\\ &\subset \left[\gamma\tilde{x}^T\ \gamma\tilde{x}_f^T\ \gamma\eta^T\ r^T\ 2P^{\frac{1}{2}}\ 2Q^{\frac{1}{2}}\right]\\ &\quad K\left[\dot{\tilde{x}}^T\ \dot{\tilde{x}}_f^T\ \dot{\eta}^T\ \dot{r}^T\ \frac{1}{2}P^{-\frac{1}{2}}\dot{P}\ \frac{1}{2}Q^{-\frac{1}{2}}\dot{Q}\right]^T.\end{aligned}$$

After substituting the dynamics from (5), (7)-(9), (12), (22), (23) and (25) and adding and subtracting $\alpha(\tilde{x}+\tilde{x}_f)^T\hat{N}_2$ and using (15), $\dot{\tilde{V}}(y)$ can be rewritten as

$$\begin{aligned}\dot{\tilde{V}} \le\ & \gamma\tilde{x}^T(r-\alpha\tilde{x}-\eta)+\gamma\tilde{x}_f^T(\eta-\alpha\tilde{x}_f)\\ &+\gamma\eta^T[-(k+\alpha)r-\alpha\eta+\tilde{x}-\tilde{x}_f]-\alpha(\tilde{x}+\tilde{x}_f)^T\hat{N}_2\\ &+\alpha(\tilde{x}+\tilde{x}_f)^T\{\tilde{W}_f^T\hat{\sigma}_f+\hat{W}_f^T\hat{\sigma}'_f[\tilde{V}_{f_1}^T\hat{x}_1+\tilde{V}_{f_2}^T\hat{x}_2]\\ &+\sum_{i=1}^m\tilde{W}_{gi}^T\hat{\sigma}_{gi}u_i+\sum_{i=1}^m\hat{W}_{gi}^T\hat{\sigma}'_{gi}[\tilde{V}_{gi_1}^T\hat{x}_1+\tilde{V}_{gi_2}^T\hat{x}_2]u_i\}\\ &+r^T[\tilde{N}+N-kr-\beta_1sgn(\tilde{x}+\tilde{x}_f)+\gamma(k+\alpha)\eta]\\ &-\gamma r^T\tilde{x}-r^T(N_1-\beta_1sgn(\tilde{x}+\tilde{x}_f))\\ &-(\dot{\tilde{x}}+\dot{\tilde{x}}_f)^TN_2+\sqrt{2}\rho(\|z\|)\|z\|^2\\ &-\alpha tr(\tilde{W}_f^T\Gamma_{wf}^{-1}\dot{\hat{W}}_f)-\alpha tr(\tilde{V}_{f_1}^T\Gamma_{vf_1}^{-1}\dot{\hat{V}}_{f_1})\\ &-\alpha tr(\tilde{V}_{f_2}^T\Gamma_{vf_2}^{-1}\dot{\hat{V}}_{f_2})-\alpha\sum_{i=1}^m tr(\tilde{W}_{gi}^T\Gamma_{wg}^{-1}\dot{\hat{W}}_{gi})\\ &-\alpha\sum_{i=1}^m tr(\tilde{V}_{gi_1}^T\Gamma_{vgi_1}^{-1}\dot{\hat{V}}_{gi_1})-\alpha\sum_{i=1}^m tr(\tilde{V}_{gi_2}^T\Gamma_{vgi_2}^{-1}\dot{\hat{V}}_{gi_2}),\end{aligned}$$

where the fact that $(r^T-r^T)_iSGN(\tilde{x}_i+\tilde{x}_{f_i})=0$ is used (the subscript i denotes the i^{th} element), where $K[sgn(\tilde{x}+\tilde{x}_f)]=SGN(\tilde{x}+\tilde{x}_f)$ [29], such that $SGN(\tilde{x}_i+\tilde{x}_{f_i})=1$ if $(\tilde{x}_i+\tilde{x}_{f_i})>0$, $[-1,1]$ if $(\tilde{x}_i+\tilde{x}_{f_i})=0$, and -1 if $(\tilde{x}_i+\tilde{x}_{f_i})<0$. Substituting the weight update laws in (17) and cancelling common terms, the above expression can be upper bounded as

$$\begin{aligned}\dot{\tilde{V}} \le\ & -\alpha\gamma\tilde{x}^T\tilde{x}-\alpha\gamma\tilde{x}_f^T\tilde{x}_f-\alpha\gamma\eta^T\eta-kr^Tr\\ &+\alpha(\tilde{x}+\tilde{x}_f)^T\tilde{N}_2+r^T\tilde{N}+\sqrt{2}\rho(\|z\|)\|z\|^2.\end{aligned} \tag{30}$$

Using (18), (20), the fact that

$$\alpha\zeta_5\|\tilde{x}+\tilde{x}_f\|\|z\|\le\alpha^2\zeta_5^2\|\tilde{x}\|^2+\alpha^2\zeta_5^2\|\tilde{x}_f\|^2+\frac{1}{2}\|z\|^2,$$

substituting $k=k_1+k_2$, and completing the squares, the expression in (30) can be further bounded as

$$\begin{aligned}\dot{\tilde{V}} \le\ & -\alpha(\gamma-\alpha\zeta_5^2)\|\tilde{x}\|^2-\alpha(\gamma-\alpha\zeta_5^2)\|\tilde{x}_f\|^2-\alpha\gamma\|\eta\|^2\\ &-k_1\|r\|^2+\left(\frac{1}{2}+\frac{\zeta_1^2}{4k_2}+\sqrt{2}\rho(\|z\|)\right)\|z\|^2.\end{aligned}$$

Provided the sufficient conditions in (26) are satisfied, the above expression can be rewritten as

$$\dot{\tilde{V}}\le-\sqrt{2}(\lambda-\frac{\zeta_1^2}{4\sqrt{2}k_2}-\rho(\|z\|))\|z\|^2\le-U(y)\quad\forall y\in\mathcal{D}, \tag{31}$$

where λ is defined in (27) and $U(y)=c\|z\|^2$, for some positive constant c, is a continuous positive semi-definite function which is defined on the domain

$$\mathcal{D}\triangleq\left\{y(t)\in R^{4n+2}|\ \|y(t)\|\le\rho^{-1}(\lambda-\frac{\zeta_1^2}{4\sqrt{2}k_2})\right\}.$$

The size of the domain $\mathcal{D}$ can be increased by increasing the gains k and α. The inequalities in (29) and (31) show that $V(y)\in\mathcal{L}_\infty$ in the domain $\mathcal{D}$; hence,

$\tilde{x}(t), \tilde{x}_f(t), \eta(t), r(t), P(t)$ and $Q(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$; (5)-(9) are used to show that $\dot{\tilde{x}}(t), \dot{\tilde{x}}_f(t), \dot{\eta}(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$. Since $x_1(t), x_2(t) \in \mathcal{L}_\infty$ by Assumption 1, $\hat{x}_1(t), \hat{x}_2(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$ using (4). Since $\tilde{x}(t), \tilde{x}_f(t), \eta(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$, using (11), $v(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$. Since $W_f, W_{gi}, \sigma_f(\cdot), \sigma_{gi}(\cdot), \varepsilon_f(\cdot), \varepsilon_{gi}(\cdot) \in \mathcal{L}_\infty, i = 1...m$, by Assumption 5-7, the control input $u(t)$ and the disturbance $d(t)$ are bounded by Assumption 2-3, and $\hat{W}_f(t), \hat{W}_{gi}(t) \in \mathcal{L}_\infty, i = 1...m$, by the use of the $proj(\cdot)$ algorithm, from (10), $\dot{r}(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$; then $\dot{z}(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$, by using (19). Hence, $U(y)$ is uniformly continuous in $\mathcal{D}$.

Let $S \subset \mathcal{D}$ denote a set defined as

$$S \triangleq \left\{ y(t) \in \mathcal{D} | U_2(y(t)) < \varepsilon_1 (\rho^{-1}(\lambda - \frac{\zeta_1^2}{4\sqrt{2}k_2}))^2 \right\}. \quad (32)$$

The region of attraction in (32) can be made arbitrarily large to include any initial condition by increasing the control gains k and α (i.e. a semi-global type of stability result), and hence

$$c \left\| z \right\|^2 \to 0 \quad \text{as } t \to \infty \quad \forall y(0) \in \mathcal{S},$$

and using the definition of $z(t)$ the following result can be proven

$$\left\| \tilde{x}(t) \right\|, \left\| \eta(t) \right\|, \left\| r(t) \right\| \to 0 \text{ as } t \to \infty \quad \forall y(0) \in \mathcal{S}.$$

From (5), it can be further shown that

$$\left\| \dot{\tilde{x}}(t) \right\| \to 0 \text{ as } t \to \infty \quad \forall y(0) \in \mathcal{S}.$$

■

IV. Conclusion

The novel design of an adaptive observer using dynamic neural networks for uncertain second-order nonlinear systems is proposed. The DNN works in conjunction with a dynamic filter without any off-line training phase. A sliding feedback term is added to the DNN structure to account for reconstruction errors and external disturbances. The observation states are proven to asymptotically converge to the system states through Lyapunov stability analysis.

References

[1] S. Diop, J. W. Grizzle, P. E. Moraal, and A. G. Stefanopoulou, "Interpolation and numerical differentiation for observer design," in *Proc. American Control Conf.*, 1994, pp. 1329–1333.

[2] S. Diop, J. W. Grizzle, and F. Chaplais, "On numerical differentiation algorithms for nonlinear estimation," in *Proc. IEEE Conf. Decis. Control*, 2000, pp. 1133–1138.

[3] J.-J. Slotine, J. Hedrick, and E. A. Misawa, "On sliding observers for nonlinear systems," in *American Control Conference*, 1986, pp. 1794–1800.

[4] C. Canudas De Wit and J.-J. Slotine, "Sliding observers for robot manipulators," *Automatica*, vol. 27, pp. 859–864, 1991.

[5] M. Mohamed, T. Karim, and B. Safya, "Sliding mode observer for nonlinear mechanical systems subject to nonsmooth impacts," in *Int. Multi-Conf. Syst. Signals Devices*, 2010.

[6] K. W. Lee and H. K. Khalil, "Adaptive output feedback control of robot manipulators using high-gain observer," *INT. J. CONTROL*, vol. 67, pp. 869–886, 1997.

[7] E. Shin and K. Lee, "Robust output feedback control of robot manipulators using high-gain observer," in *IEEE International Conference on Control and Applications*, 1999.

[8] A. Astolfi, R. Ortega, and A. Venkatraman, "A globally exponentially convergent immersion and invariance speed observer for n degrees of freedom mechanical systems," in *Proc. Chin. Control Conf.*, 2009.

[9] N. Lotfi and M. Namvar, "Global adaptive estimation of joint velocities in robotic manipulators," *IET Control Theory Appl.*, vol. 4, pp. 2672–2681, 2010.

[10] J. Davila, L. Fridman, and A. Levant, "Second-order sliding-mode observer for mechanical systems," *IEEE Trans. Autom. Control*, vol. 50, pp. 1785–1789, 2005.

[11] D. Dawson, Z.Qu, and J. Carroll, "On the state observation and output feedback problems for nonlinear uncertain dynamic systems," *Syst. Contr. Lett.*, vol. 18, pp. 217–222, 1992.

[12] B. Xian, M. S. de Queiroz, D. M. Dawson, and M. McIntyre, "A discontinuous output feedback controller and velocity observer for nonlinear mechanical systems," *Automatica*, vol. 40, no. 4, pp. 695–700, 2004.

[13] Y. H. Kim and F. L. Lewis, "Neural network output feedback control of robot manipulators," *IEEE Trans. Robot. Autom.*, vol. 15, pp. 301–309, 1999.

[14] M. Polycarpou and P. Ioannou, "Identification and control of nonlinear systems using neural network models: Design and stability analysis," *Systems Report 91-09-01, University of Southern California*, 1991.

[15] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Math. Control Signals Syst.*, vol. 2, pp. 303–314, 1989.

[16] A. Barron, "Universal approximation bounds for superpositions of a sigmoidal function," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 930–945, 1993.

[17] F. L. Lewis, R. Selmic, and J. Campos, *Neuro-Fuzzy Control of Industrial Systems with Actuator Nonlinearities*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2002.

[18] W. E. Dixon, A. Behal, D. M. Dawson, and S. Nagarkatti, *Nonlinear Control of Engineering Systems: A Lyapunov-Based Approach*. Birkhäuser Boston, 2003.

[19] M. Krstic, P. V. Kokotovic, and I. Kanellakopoulos, *Nonlinear and Adaptive Control Design*, S. Haykin, Ed. John Wiley & Sons, 1995.

[20] K. Funahashi and Y. Nakamura, "Approximation of dynamic systems by continuous-time recurrent neural networks," *Neural Networks*, vol. 6, pp. 801–806, 1993.

[21] B. Xian, D. M. Dawson, M. S. de Queiroz, and J. Chen, "A continuous asymptotic tracking control strategy for uncertain nonlinear systems," *IEEE Trans. Autom. Control*, vol. 49, no. 7, pp. 1206–1211, Jul. 2004.

[22] H. Dinh, S. Bhasin, and W. E. Dixon, "Dynamic neural network-based robust identification and control of a class of nonlinear systems," in *Proc. IEEE Conf. Decis. Control*, Dec. 2010, pp. 5536–5541.

[23] A. Filippov, "Differential equations with discontinuous right-hand side," *Am. Math. Soc. Transl.*, vol. 42, pp. 199–231, 1964.

[24] A. F. Fillipov, *Differential Equations with Discontinuous Right-hand Sides*. Kluwer Academic Publishers, 1988.

[25] G. V. Smirnov, *Introduction to the theory of differential inclusions*. American Mathematical Society, 2002.

[26] J. P. Aubin and H. Frankowska, *Set-valued analysis*. Birkhäuser, 2008.

[27] F. H. Clarke, *Optimization and nonsmooth analysis*. SIAM, 1990.

[28] D. Shevitz and B. Paden, "Lyapunov stability theory of nonsmooth systems," *IEEE Trans. Autom. Control*, vol. 39, pp. 1910–1914, 1994.

[29] B. Paden and S. Sastry, "A calculus for computing Filippov's differential inclusion with application to the variable structure control of robot manipulators," *IEEE Trans. Circuits Syst.*, vol. 34, pp. 73–82, 1987.