Extended state observer for MIMO nonlinear systems with stochastic uncertainties

Ze-Hao Wu and Bao-Zhu Guo

School of Mathematics and Big Data, Foshan University, Foshan, Guangdong 528000, People’s Republic of China

ABSTRACT
In this paper, both linear extended state observer (ESO) and nonlinear ESO with homogeneous weighted functions are proposed for a class of multi-input multi-output (MIMO) nonlinear systems composed of coupled subsystems with large stochastic uncertainties. The stochastic uncertainties in each subsystem including internal coupled unmodelled dynamics and external stochastic disturbance without known statistical characteristics are lumped together as the stochastic total disturbance (extended state) of each subsystem. The linear ESO and nonlinear ESO are designed separately for real-time estimation of not only the unmeasured state but also the stochastic total disturbance of each subsystem. The practical mean square convergence of these two classes of ESOs are developed. Some numerical simulations are presented to demonstrate the effectiveness of the ESOs with the advantages of smaller peaking values and more accurate estimation by the nonlinear ESO.

1. Introduction
The active disturbance rejection control (ADRC), as an emerging control technology in dealing with vast uncertainties, was first proposed by Jingqing Han as an alternative of proportional-integral-derivative control in the late 1980s (Han, 2009). In the past two decades, ADRC has been successfully applied to many engineering control problems such as synchronous motors (Sira-Ramirez, Linares-Flores, Garcia-Rodriguez, & Contreras-Ordaz, 2014), DC-DC power converter (Sun & Gao, 2005), control system in superconducting radio frequency cavities (Vincent et al., 2011), flight vehicles control (Xia & Fu, 2013), and gasoline engines (Xue et al., 2015), among many others. In particular, ADRC has been hardwired into new motor control chips made by industry giants such as Texas Instruments (Technical Reference Manual, 2013) and Freescale Semiconductor (Kinetics Motor Suite, 2015) and implemented successfully by the utility industry in power plants (Sun, Li, Hu, Lee, & Pan, 2016). It has been reported in Zheng and Gao (2012) that ADRC has been tested in Parker Hannifin Parflex hose extrusion plant and across multiple production lines for over 8 months with the product performance capability index (Cpk) improved by 30% and the energy consumption reduced over 50%.

Extended state observer (ESO) is the key part of ADRC proposed in Han (1995), which is used not only to estimate in real time the unmeasured state but also the ‘total disturbance’ which may come from unmodelled system dynamics, unknown coefficient of control, and external disturbance, or even if whatever the part of hardly to be dealt with by the practitioner. Based on the estimation of total disturbance and unmeasured state of system by ESO, we can design an output feedback control which is almost free of mathematical models to cancel (compensate) the total disturbance in the feedback loop and thus obtain desired control performance. This remarkable feature of ESO makes ADRC a very different nature in dealing with vast uncertainties. Moreover, as a significant breakthrough in observer design, ESO is an extension of traditional state observer where only the unmeasured state of system is estimated.

Roughly speaking, ESO deals with systems with uncertainties coming from either the system itself or from the external disturbance. In this unconventional idea, all the uncertainties influencing the performance of system are lumped together into ‘total disturbance’ or ‘extended state’ and is then estimated by ESO. The first ESO which uses nonlinear gains was designed in Han (1995) as

\[
\begin{align*}
\dot{x}_1(t) &= \dot{x}_2(t) - \alpha_1 g_1(\dot{x}_1(t) - y(t)), \\
\dot{x}_2(t) &= \dot{x}_3(t) - \alpha_2 g_2(\dot{x}_1(t) - y(t)), \\
&\vdots \\
\dot{x}_n(t) &= \dot{x}_{n+1}(t) - \alpha_n g_n(\dot{x}_1(t) - y(t)) + u(t), \\
\dot{x}_{n+1}(t) &= -\alpha_{n+1} g_{n+1}(\dot{x}_1(t) - y(t)),
\end{align*}
\]

(1)

which is for a general n-dimensional single-input single-output (SISO) system as follows:

\[
x^{(n)}(t) = f(t, x(t), \dot{x}(t), \ldots, x^{(n-1)}(t)) + w(t) + u(t),
\]

\[
y(t) = x(t)
\]
that can be rewritten as
\[
\begin{align*}
\dot{x}_1(t) &= x_2(t), \\
\dot{x}_2(t) &= x_3(t), \\
&\vdots \\
\dot{x}_n(t) &= f(t, x_1(t), \ldots, x_n(t)) + w(t) + u(t), \\
y(t) &= x_1(t),
\end{align*}
\]
where \(u(t)\) is the control input, \(y(t)\) is the measured output, \(f(\cdot) \in C(\mathbb{R}^n)\) is the unknown system function, and \(w(t)\) is the external disturbance. The \(x_{n+1}(t) = f(t, x_1(t), \ldots, x_n(t)) + w(t)\), as a signal of time, represents the total effect of both the nonlinear, time-varying, and unmodelled dynamics of the plant and the external disturbance and is regarded as the total disturbance (extended state). The main idea of ESO (1) is that for some appropriately chosen functions \(g_i(\cdot) \in C(\mathbb{R})\), the states \(\hat{x}_i(t) (i = 1, 2, \ldots, n)\) and \(\hat{x}_{n+1}(t)\) of the observer can approximately recover the states \(x_i(t) (i = 1, 2, \ldots, n)\) and the total disturbance (extended state) of system (2), respectively, by regulating constants \(\alpha_i\), i.e. \(\hat{x}_i(t) \approx x_i(t) (i = 1, 2, \ldots, n + 1)\).

The multiple choice of tuning parameters of ESO (1) has been replaced with the one-parameter tuning linear ESO proposed in Gao (2003) in terms of bandwidth as follows:
\[
\begin{align*}
\dot{x}_1(t) &= \hat{x}_2(t) + \frac{k_1}{\varepsilon} (y(t) - \hat{x}_1(t)), \\
\dot{x}_2(t) &= \hat{x}_3(t) + \frac{k_2}{\varepsilon^2} (y(t) - \hat{x}_1(t)), \\
&\vdots \\
\dot{x}_n(t) &= \hat{x}_{n+1}(t) + \frac{k_n}{\varepsilon^{n-1}} (y(t) - \hat{x}_1(t)) + u(t), \\
\dot{\hat{x}}_{n+1}(t) &= \frac{k_{n+1}}{\varepsilon^n} (y(t) - \hat{x}_1(t)),
\end{align*}
\]
where \(k_i (i = 1, 2, \ldots, n + 1)\) are designed parameters such that the following matrix is Hurwitz:
\[
E = \begin{pmatrix}
-k_1 & 1 & 0 & \cdots & 0 \\
-k_2 & 0 & 1 & \cdots & 0 \\
& \cdots & \cdots & \cdots & \cdots \\
-k_n & 0 & 0 & \cdots & 1 \\
-k_{n+1} & 0 & 0 & \cdots & \cdots \\
\end{pmatrix}_{(n+1) \times (n+1)}
\]
and \(\varepsilon > 0\) is the tuning parameter. Here and throughout the paper, we always drop \(\varepsilon\) for the solution of (3) by abuse of notation without confusion. The convergence of linear ESO (3) for uncertain SISO systems was presented in Shao and Gao (2017) and Zheng, Gao, and Gao (2007).

Although linear ESO is easy to design in practices, a nonlinear ESO (5) which is called homogeneous ESO in literatures has been shown to be more accurate with small peaking values in estimating unmeasured state and total disturbance compared with the linear ESO, see, for instance (Guo & Zhao, 2011, 2012, 2016; Zhao & Guo, 2015). As a special case of ESO (1) and a nonlinear extension of linear ESO (3), the homogeneous ESO is a one-parameter tuning nonlinear ESO with homogeneous weighted functions as follows:
\[
\begin{align*}
\dot{x}_1(t) &= \hat{x}_2(t) + \varepsilon^{n-1} k_1 \left[ \frac{y(t) - \hat{x}_1(t)}{\varepsilon^n} \right]^a, \\
\dot{x}_2(t) &= \hat{x}_3(t) + \varepsilon^{n-2} k_2 \left[ \frac{y(t) - \hat{x}_1(t)}{\varepsilon^n} \right]^{2a-1}, \\
&\vdots \\
\dot{x}_n(t) &= \hat{x}_{n+1}(t) + k_n \left[ \frac{y(t) - \hat{x}_1(t)}{\varepsilon^n} \right]^{na-(n-1)} + u(t), \\
\dot{\hat{x}}_{n+1}(t) &= \frac{1}{\varepsilon} k_{n+1} \left[ \frac{y(t) - \hat{x}_1(t)}{\varepsilon^n} \right]^{(n+1)a-n},
\end{align*}
\]
where \(\varepsilon\) is the tuning parameter, and \(|\theta|^s = \text{sign}(\theta)|\theta|^s\) for all \(\theta \in \mathbb{R}\).

Motivated from the convergence of homogeneous state observer for systems without uncertainty investigated in Menard, Moulay, and Perruquetti (2010) and Shen and Xia (2008), the convergence of homogeneous ESO (5) for uncertain SISO nonlinear systems, uncertain MIMO nonlinear systems, and uncertain lower triangular nonlinear systems have been proved in Guo and Zhao (2011, 2012) and Zhao and Guo (2015), respectively. Very recently, the monograph (Guo & Zhao, 2016) presented a comprehensive theoretical foundations of homogeneous ESO for uncertain nonlinear systems.

However, most of the available literatures mainly address ESO for uncertain systems without stochastic characteristics. For the state estimation problem of nonlinear stochastic systems driven by Brownian motions, we could find some breakthrough efforts like (Tarn & Rasis, 1976; Wang & Zhang, 2012; Yaz & Azemi, 1993), and the adaptive state and parameter estimators were proposed simultaneously for the stochastic systems with unknown parameters (Xie & Khargonekar, 2012). But very few have been done for the state estimation problem of systems with vast stochastic uncertainties including unmodelled dynamics and external stochastic disturbance without known statistical characteristics. We could find some efforts like the applications of fuzzy logic systems in approximating the unknown nonlinear system functions and the fuzzy state observer in estimating the unmeasured states of uncertain nonlinear systems (Tong, Li, & Sui, 2016; Tong, Zhang, & Li, 2016) and stochastic nonlinear systems (Li, Sui, & Tong, 2017; Li, Tong, & Li, 2015). Very recently, ESO has been adopted for the practical mean square estimation of both unmeasured state and stochastic total disturbance that includes internal unknown dynamics and external stochastic disturbance without the known statistical characteristics for SISO systems (Guo, Wu, & Zhou, 2016; Wu & Guo, 2017). Along the line of (Guo et al., 2016; Wu & Guo, 2017), in this paper, we generalise both the linear ESO and homogeneous ESO to the MIMO nonlinear systems with vast stochastic uncertainties as follows:
\[
x_1^{(m)}(t) = f_1(x_1(t), \ldots, x_1^{(m-1)}(t), \ldots, x_m^{(m-1)}(t), w_1(t)) + h_1(u(t)),
\]
\[ x_2^{(n_2)}(t) = f_2(x_1(t), \ldots, x_1^{(m-1)}(t), \ldots, x_m^{(m-1)}(t), w_2(t)) + h_2(u(t)), \]
\[ \vdots \]
\[ x_m^{(n_m)}(t) = f_m(x_1(t), \ldots, x_1^{(m-1)}(t), \ldots, x_m^{(m-1)}(t), w_m(t)) + h_m(u(t)), \]
\[ y_i(t) = x_i(t), \quad i = 1, 2, \ldots, m \]  
(6)

that can be rewritten as
\[ \dot{x}_{i,1}(t) = x_{i,2}(t), \]
\[ \dot{x}_{i,2}(t) = x_{i,3}(t), \]
\[ \vdots \]
\[ \dot{x}_{i,n_i}(t) = f_i(x_{i,1}(t), \ldots, x_{i,n_i}(t), \ldots, x_{m,n_m}(t), w_i(t)) + h_i(u(t)), \]
\[ y_i(t) = x_{i,1}(t), \quad i = 1, 2, \ldots, m, \]  
(7)

where \( x(t) = (x_1(t), \ldots, x_m(t)) \in \mathbb{R}^m \) is the state with \( x_i(t) = (x_{i,1}(t), \ldots, x_{i,n_i}(t)) \in \mathbb{R}^{n_i} \) and \( n = n_1 + \cdots + n_m, \ u(t) = (u_1(t), \ldots, u_m(t)) \in \mathbb{R}^m \) is the input, \( y(t) = (y_1(t), \ldots, y_m(t)) \in \mathbb{R}^m \) is the output, \( w_i(t) \in \mathbb{R} \) is the external stochastic disturbance, respectively; \( i = 1, 2, \ldots, m \), and
\[ x_{i,n_i+1}(t) \triangleq f_i(x_{i,1}(t), \ldots, x_{i,n_i}(t), \ldots, x_{m,n_m}(t), w_i(t)) = f_i(x(t), w_i(t)) \]  
(8)

is the stochastic total disturbance (extended state) in \( i \)-th subsystem representing the combined effect of internal coupled unmodelled dynamics and external stochastic disturbance, where \( f_i(\cdot) \) is continuously differentiable and twice continuously differentiable with respect to \( x \) and \( w_i \), respectively. The \( w_i(t) = \psi_i(t, B_i(t)) \in \mathbb{R} \) for some bounded unknown function \( \psi_i(\cdot): [0, \infty) \times \mathbb{R} \rightarrow \mathbb{R} \) is the external stochastic disturbance, where \( \{B_i(t)\}_{t \geq 0} \) is the standard one-dimensional Brownian motion defined on a complete probability space \( (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P) \) with \( \Omega \) being a sample space, \( \mathcal{F} \) a \( \sigma \)-field, \( \{\mathcal{F}_t\}_{t \geq 0} \) a filtration, and \( P \) the probability measure.

It should be noted that system (7) and the external stochastic disturbance \( w_i(t) \) here are quite general. First, the statistical characteristic of \( w_i(t) \) is not necessarily to be known since the function \( \psi_i(\cdot) \) is unknown. Second, the \( w_i(t) \) covers the disturbance investigated via ESO in aforementioned literatures when \( \psi_i(\cdot) \) is the functions of time variable \( t \) only: \( w(t) \triangleq \psi_i(t) \). In this case, system (7) covers some special systems studied in existing literatures such as SISO systems in Guo and Zhao (2011) and MIMO systems in Guo and Zhao (2012). In addition, for the stochastic case, the bounded stochastic noise considered in Huang, Zhu, Ni, and Ko (2002), Hu, Chen, and Zhu (2012), Huang and Zhu (2004) and Li, Xu, Yang, and Sun (2008) is also covered. Finally, system (7) covers special SISO nonlinear systems with stochastic disturbance investigated via ESO in Guo et al. (2016) as a special case of \( k = m = 1 \).

The main contributions of this paper are: (a) A novel estimation strategy ESO is systematically proposed for a class of MIMO nonlinear systems with stochastic uncertainties to estimate not only the unmeasured state but also the stochastic total disturbance in large scale, whereas most of the available literature for nonlinear stochastic systems address only on state estimation. (b) Our stochastic uncertainties are very large, which can include internal coupled unmodelled dynamics and external stochastic disturbance without statistical characteristics. In other words, ESO is an almost model-free estimation approach. (c) Nonlinear ESO with homogeneous weighted functions are first designed for MIMO nonlinear systems with vast stochastic uncertainties, by which the smaller peaking values and more accurate estimation compared with linear ESO are observed and the convergence is presented completely.

We proceed as follows. In the next section, Section 2, we give some preliminaries about finite-time stability and homogeneity for the proof of convergence of homogeneous ESO. Both linear ESO and homogeneous ESO are designed for system (7), and practical mean square convergence for these two classes of ESOs are stated in Section 3. The proofs of the main results are presented in Section 4. In Section 5, some numerical experiments are performed to illustrate effectiveness of the ESOs, and some comparisons between linear ESO and homogeneous ESO are made numerically.

### 2. Preliminaries

The following notations are used throughout the paper. The \( \mathbb{R}^s \) represents the \( s \)-dimensional Euclidean space; \( C(\mathbb{R}^s, \mathbb{R}) \) stands for the space of all continuous functions from \( \mathbb{R}^s \) into \( \mathbb{R} \); For a vector or matrix \( X, X^T \) denotes its transpose; \( I_{s \times s} \) denotes the \( s \times s \) unit matrix; \( \lambda_{\min}(X) \) and \( \lambda_{\max}(X) \) represent the minimal and maximal eigenvalues of the symmetric real matrix \( X \), respectively; \( \|X\| \) denotes the Euclidean norm of the vector \( X \). For a differentiable function \( f: \mathbb{R}^s \rightarrow \mathbb{R}, df/\partial \vartheta \triangleq (df/\partial \vartheta_1, \ldots, df/\partial \vartheta_j) \) for \( \vartheta = (\vartheta_1, \ldots, \vartheta_j) \in \mathbb{R}^j; \) \( |\vartheta|^r \triangleq \text{sign}(\vartheta) |\vartheta|^r \) for all \( \vartheta \in \mathbb{R}^j \).

To deal with the convergence of homogeneous ESO, finite-time stability and homogeneity are introduced as follows.

**Definition 2.1:** The zero equilibrium of the following system
\[ \dot{z}(t) = f(z(t)), \quad z(0) = z_0 \in \mathbb{R}^s \]  
(9)

is said to be globally finite-time stable, if the zero equilibrium of system (9) is Lyapunov stable and for any \( z_0 \in \mathbb{R}^s \), there exists \( T(z_0) > 0 \) such that the solution of (9) satisfies that
\[ \lim_{t \rightarrow T(z_0)} z(t) = 0 \text{ and } z(t) = 0 \text{ for all } t \in [T(z_0), \infty). \]

**Definition 2.2:** A function \( f: \mathbb{R}^s \rightarrow \mathbb{R} \) is said to be homogeneous of degree \( d \) with respect to weights \( \{r_j > 0\}_{j=1}^s \), if
\[ V(\lambda^{r_1} z_1, \lambda^{r_2} z_2, \ldots, \lambda^{r_s} z_s) = \lambda^d V(z_1, z_2, \ldots, z_s) \]  
(10)

for all \( \lambda > 0 \) and all \( (z_1, z_2, \ldots, z_s) \in \mathbb{R}^s \).

A vector field \( f: \mathbb{R}^s \rightarrow \mathbb{R} \) is said to be homogeneous of degree \( d \) with respect to weights \( \{r_j > 0\}_{j=1}^s \), if for all \( j = 1, 2, \ldots, s \), the \( j \)-th component \( f_j \) is a homogeneous function of
degree $d + r_j$, that is,

$$f_j(\lambda^1 z_1, \lambda^2 z_2, \ldots, \lambda^r z_r) = \lambda^{d+r} f_j(z_1, z_2, \ldots, z_r)$$  \hspace{1cm} (11)

for all $\lambda > 0$ and all $(z_1, z_2, \ldots, z_r) \in \mathbb{R}^r$. The system (9) is homogeneous of degree $d$ if the vector field $f$ is homogeneous of degree $d$.

The following Lemma 2.1 comes from Theorem 2 of Rosier (1992) and Theorem 6.2 of Bhat and Bernstein (2005).

**Lemma 2.1:** If system (2.1) is homogeneous of degree $d$ with weights $\{r_j\}_{j=1}^n$, and its zero equilibrium is globally asymptotically stable, then for any $\theta > \max_{1 \leq j \leq s} (-d, r_j)$, there exists a positive definite, radially unbounded, differentiable function $V : \mathbb{R}^s \rightarrow \mathbb{R}$ such that $V$ is homogeneous of degree $\theta$ with respect to weights $\{r_j\}_{j=1}^n$, and the Lie derivative of $V(z)$ along the vector field $f$: $L_f V(z) = \langle \nabla V(z), f(z) \rangle$ is negative definite.

The following Lemma 2.2 is the Lemma 4.2 of Bhat and Bernstein (2005).

**Lemma 2.2:** Let $V_1, V_2 : \mathbb{R}^r \rightarrow \mathbb{R}$ be continuous functions, homogeneous of degree $l_1 > 0, l_2 > 0$ with respect to the same weights, respectively, and $V_1$ is positive definite. Then for each $z \in \mathbb{R}^r$,

$$\min_{\theta \in V_1^{-1}(1)} V_2(\theta) \left( V_1(z) \right)^{l_2/l_1} \leq V_2(z) \leq \max_{\theta \in V_1^{-1}(1)} V_2(\theta) \left( V_1(z) \right)^{l_2/l_1},$$  \hspace{1cm} (12)

where $V_1^{-1}(1) = \{ \theta \in \mathbb{R}^m | V_1(\theta) = 1 \}$.

The Lemmas 2.3–2.4 come from Perruquetti, Floquet, and Moulay (2008) directly.

**Lemma 2.3:** If $a_i \in (1 - (1/n_i), 1)$, then the vector field

$$\Phi_i(\theta) \triangleq (\Phi_i,1(\theta), \ldots, \Phi_i,n_i(\theta), \Phi_i,n_i+1(\theta))$$

$$\triangleq (\partial_2 - k_{i,1} \partial_1 a_i, \ldots, \partial_{n_i} - k_{i,1} \partial_1 n_i a_i - (n_i - 1),$$

$$- k_{i,n_i+1} \partial_1 n_i a_i - n_i, \forall \theta = (\partial_1, \ldots, \partial_{n_i+1}) \in \mathbb{R}^{n_i+1}$$  \hspace{1cm} (13)

is homogeneous of degree $-d_i = a_i - 1$ with respect to weights $\{r_{ij} = (j - 1) a_i - (j - 2) \}_{j=1}^{n_i+1}$.

**Lemma 2.4:** There exists $a_i^* \in (1 - (1/n_i), 1)$ such that for any $a_i \in (a_i^*, 1)$, if the matrix $E_i$ in (16) is Hurwitz, then the system $\dot{\theta}(t) = \Phi_i(\theta(t))$ is globally finite-time stable.

3. ESO design and main results

In this section, we design an estimator in terms of the input $u(t)$ and the measured output $y_i(t)$ only, to estimate both the unmeasured state and the stochastic total disturbance of each subsystem. To this end, we need the following assumptions.

Assumption 3.1 is about the unknown function $\psi_i(\cdot)$ that defines the external stochastic disturbance.

**Assumption 3.1:** The $\psi_i(t, \theta) : [0, \infty) \times \mathbb{R} \rightarrow \mathbb{R}$ is continuously differentiable and twice continuously differentiable with respect to $t$ and $\theta$, respectively, and there exist some (known) constant $C_{1,i} > 0$ such that for all $t \geq 0, \theta \in \mathbb{R}$,

$$|\psi_i(t, \theta)| + \left| \frac{\partial \psi_i(t, \theta)}{\partial t} \right| + \left| \frac{\partial^2 \psi_i(t, \theta)}{\partial \theta^2} \right| \leq C_{1,i}.$$  \hspace{1cm} (14)

The Assumption 3.2 is a prior assumption about the input $u(t)$ and the solution $x(t)$ of (7).

**Assumption 3.2:** There exists a (known) constant $M \geq 0$ such that $\|u(t)\| + \|x(t)\| \leq M$ almost surely for all $t \geq 0$.

**Remark 3.1:** Since the external stochastic disturbance $w_i(t) = \psi_i(t, B_i(t)) \in \mathbb{R}$ is coupled with unmodelled dynamics to be estimated by ESO, it is reasonable to assume that the $\psi_i(\cdot)$ itself and its first-order and the second-order derivatives with respect to its arguments are bounded, i.e. the external stochastic disturbance itself and its ‘variation’ are confined to be bounded. It is also easy to verify that many practical disturbances such as $\cos(a + bt)$, $\sin(a + bt)$ and stochastic disturbances such as $\cos(at + bB_i(t))$, $\sin(at + bB_i(t))$ satisfy the assumption, where $a, b$ are constants and the former disturbances are covered by letting $\psi_i(t)$ be the function with respect to the time variable only.

**Remark 3.2:** It should be noticed that the boundedness of the state in Assumption 3.2 exists widely in most practical control systems such as those for faults diagnosis (Yan, Tian, Shi, & Weng, 2008). Since we are only concerned with the practical mean square convergence of both linear ESO and nonlinear ESO for open-loop system, the boundedness of state in Assumption 3.2 is used for estimation of state-dependent stochastic total disturbance (8). If the estimated stochastic total disturbance is state-independent, the boundedness of the state in Assumption 3.2 can be removed because the boundedness assumption is used only in deduction of (31) in the proof of Theorem 3.1 or (55) in the proof of Theorem 3.2. In addition, since both linear ESO and nonlinear ESO are designed for feedback purpose, we can also use feedback to make the system state bounded, which is discussed in system (72) in numerical simulations. Finally, the input must be bounded in engineering practice as assumed in Assumption 3.2. Finally, the boundedness assumption for the state is only for estimation of open-loop system. When we design ESO-based output feedback control, since the system state is controlled to be bounded, this assumption is not required for the closed-loop system.

The following block diagram shows the estimation procedure of ESO:
We introduce the one-parameter tuning linear ESO for system (7) as follows:

\[
\begin{align*}
\dot{x}_{i,1}(t) &= \dot{x}_{i,2}(t) + \frac{k_{i,1}}{\varepsilon}(y_i(t) - \hat{x}_{i,1}(t)), \\
\dot{x}_{i,2}(t) &= \dot{x}_{i,3}(t) + \frac{k_{i,2}}{\varepsilon^2}(y_i(t) - \hat{x}_{i,1}(t)), \\
&\vdots \\
\dot{x}_{i,n_i}(t) &= \dot{x}_{i,n_i+1}(t) + \frac{k_{i,n_i}}{\varepsilon^{n_i}}(y_i(t) - \hat{x}_{i,1}(t)) + h_i(u(t)), \\
\dot{x}_{i,n_i+1}(t) &= \frac{1}{\varepsilon}k_{i,n_i+1}\left[\frac{y_i(t) - \hat{x}_{i,1}(t)}{\varepsilon^{n_i}}\right]^{(n_i+1)n_i-n_i-1} + h_i(u(t)),
\end{align*}
\]

where \( y_i(t) \) is injected from system (7), \( \varepsilon > 0 \) is the tuning parameter, and \( k_{ij} \) \((j = 1,2,\ldots,n_i + 1)\) are designed parameters such that the following matrix is Hurwitz:

\[
E_i = \begin{pmatrix}
-k_{i,1} & 1 & 0 & \cdots & 0 \\
-k_{i,2} & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
-k_{i,n_i} & 0 & 0 & \cdots & 1 \\
-k_{i,n_i+1} & 0 & 0 & \cdots & 0
\end{pmatrix}^{(n_i+1)\times(n_i+1)}.
\]

Let \( H_i \) be the positive definite matrix solution satisfying the Lyapunov equation as follows:

\[
H_iE_i + E_i^TH_i = -I_{(n_i+1)\times(n_i+1)}.
\]

In what follows, the homogeneous ESO is constructed for system (7) as follows:

\[
\begin{align*}
\dot{x}_{i,1}(t) &= \hat{x}_{i,2}(t) + \varepsilon^{n_i-1}k_{i,1}\left[\frac{y_i(t) - \hat{x}_{i,1}(t)}{\varepsilon^{n_i}}\right]^{n_i}, \\
\dot{x}_{i,2}(t) &= \hat{x}_{i,3}(t) + \varepsilon^{n_i-2}k_{i,2}\left[\frac{y_i(t) - \hat{x}_{i,1}(t)}{\varepsilon^{n_i}}\right]^{2n_i-1}, \\
&\vdots \\
\dot{x}_{i,n_i}(t) &= \hat{x}_{i,n_i+1}(t) + k_{i,n_i}\left[\frac{y_i(t) - \hat{x}_{i,1}(t)}{\varepsilon^{n_i}}\right]^{(n_i+1)n_i-n_i-1} + h_i(u(t)), \\
\dot{x}_{i,n_i+1}(t) &= \frac{1}{\varepsilon}k_{i,n_i+1}\left[\frac{y_i(t) - \hat{x}_{i,1}(t)}{\varepsilon^{n_i}}\right]^{(n_i+1)n_i-n_i-1} + h_i(u(t)),
\end{align*}
\]

where \( k_{ij} \) \((j = 1,2,\ldots,n_i + 1)\) are constants such that the matrix \( E_i \) in (16) is Hurwitz.

Now, we are ready to state our main results. The practical mean square convergence of both linear ESO and homogeneous ESO is summarised in the succeeding Theorem 3.1 and Theorem 3.2, respectively.

**Theorem 3.1:** Under Assumptions 3.1, 3.2, the linear ESO (15) is practically mean square convergent in the sense that

(i) For any \( \varepsilon \in (0,1) \), there exists an \( \varepsilon \)-dependent constant \( t_{\varepsilon} > 0 \) (specified in (37)) such that for any given initial values of (7) and (15),

\[
\mathbb{E}|x_{ij}(t) - \hat{x}_{ij}(t)|^2 \leq \Gamma_i\varepsilon^{2n_i+3-2j}, \quad t \in [t_{\varepsilon}, \infty), \quad j = 1,2,\ldots,n_i + 1,
\]

where \( \Gamma_i \) is an \( \varepsilon \)-independent constant specified in (39); As a result,

\[
\lim_{t \to \infty} \mathbb{E}|x_{ij}(t) - \hat{x}_{ij}(t)|^2 \leq \Gamma_i\varepsilon^{2n_i+3-2j}, \quad j = 1,2,\ldots,n_i + 1;
\]

(ii) For each positive constant \( a > 0 \) and any given initial values of (7) and (15),

\[
\lim_{t \to 0} \mathbb{E}|x_{ij}(t) - \hat{x}_{ij}(t)|^2 = 0 \text{ uniformly in } t \in [a, \infty), \quad j = 1,2,\ldots,n_i + 1.
\]

**Theorem 3.2:** Suppose that Assumptions 3.1, 3.2 are satisfied and \( a_i \in (1 - (1/n_i + 1), 1) \). Then the homogeneous ESO (18) is practically mean square convergent in the sense that: there are a constant \( \varepsilon^* > 0 \) and an \( \varepsilon \)-dependent constant \( t_{\varepsilon} > 0 \) with \( \varepsilon \in (0, \varepsilon^*) \) such that for any given initial values of (7) and (18),

\[
\mathbb{E}|x_{ij}(t) - \hat{x}_{ij}(t)|^2 \leq \Gamma_{ij}\varepsilon^{2n_i+2+\Lambda_{ij}-2j}, \quad t \in [t_{\varepsilon}, \infty), \quad j = 1,2,\ldots,n_i + 1,
\]

where \( \Gamma_{ij} \) is an \( \varepsilon \)-independent constant and \( \Lambda_{ij} > 1 \) specified in (68) and (69), respectively; As a result,

\[
\lim_{t \to \infty} \mathbb{E}|x_{ij}(t) - \hat{x}_{ij}(t)|^2 \leq \Gamma_{ij}\varepsilon^{2n_i+2+\Lambda_{ij}-2j}, \quad j = 1,2,\ldots,n_i + 1.
\]

**Remark 3.3:** Compared the estimation error (22) by homogeneous ESO with (19) by linear ESO, we find that the power of \( \varepsilon \) in (22) being \( 2n_i + 2 + \Lambda_{ij} - 2j \) is larger than \( 2n_i + 3 - 2j \) in (19). Therefore, the estimation error of homogeneous ESO drops much rapidly than the one of linear ESO as \( \varepsilon \) decreases,
4. Proof of main results

Proof of Theorem 3.1: Set
\[ \eta_{i,j}(t) = \frac{x_{i,j}(\epsilon t) - \hat{x}_{i,j}(\epsilon t)}{\epsilon} ; j = 1, \ldots, n_i + 1, \quad 1 \leq j \leq n_i + 1 ; \]  
\[ \eta_i = (\eta_{i,1}, \eta_{i,2}, \ldots, \eta_{i,n_i + 1}) ; \]  
\[ \Psi_i(x_i) = (x_{i,2}, x_{i,3}, \ldots, f_i(x_i, w_i) + h_i(u)); \]  
\[ \Psi(x) = (\Psi_1(x_1), \Psi_2(x_2), \ldots, \Psi_m(x_m)) . \]

In terms of the Itô's formula, we can obtain that
\[ \text{d} x_{i,n_i+1}(t) = \frac{\partial f_i(x_i, w_i(t))}{\partial x} \text{d} t + \frac{\partial f_i(x_i, w_i(t))}{\partial w_i} \text{d} w_i + \frac{\partial^2 f_i(x_i, w_i(t))}{2} \frac{\partial^2 \Psi_i(x_i, B_i(t))}{\partial \theta^2} \text{d} \theta \]  
\[ + \frac{\partial f_i(x_i, w_i(t))}{\partial w_i} \frac{\partial \Psi_i(x_i, B_i(t))}{\partial \theta} \text{d} \theta \]  
\[ + \frac{\partial f_i(x_i, w_i(t))}{\partial w_i} \frac{\partial \Psi_i(x_i, B_i(t))}{\partial \theta} \text{d} B_i(t) \]  
\[ \triangleq \Delta_{1,i}(t) \text{d} t + \Delta_{2,i}(t) \text{d} B_i(t) . \]

Thus, system (7) could be rewritten as
\[ \text{d} x_{i,1}(t) = x_{i,2}(t) \text{d} t , \]  
\[ \text{d} x_{i,2}(t) = x_{i,3}(t) \text{d} t , \]  
\[ \vdots \]  
\[ \text{d} x_{i,n_i}(t) = (x_{i,n_i+1}(t) + h_i(u(t))) \text{d} t , \]  
\[ \text{d} x_{i,n_i+1}(t) = \Delta_{1,i}(t) \text{d} t + \Delta_{2,i}(t) \text{d} B_i(t) , \]  
\[ y_i(t) = x_{i,1}(t) , \quad i = 1, 2, \ldots, m. \]

Noting that \( \hat{B}_i(t) \triangleq (1/\sqrt{t})B_i(\epsilon t) \) is still standard one-dimensional Brownian motion defined on the complete probability space \((\Omega, F, (F_t), t \geq 0, P)\).

Considering (29) and (3), a direct computation shows that
\[ \eta_i(t) = (\eta_{i,1}(t), \eta_{i,2}(t), \ldots, \eta_{i,n_i + 1}(t)) \]  
\[ \text{satisfies the following Itô-type stochastic differential equation:} \]
\[ \text{d} \eta_{i,1}(t) = (\eta_{i,2}(t) - k_{i,1}\eta_{i,1}(t)) \text{d} t , \]  
\[ \text{d} \eta_{i,2}(t) = (\eta_{i,3}(t) - k_{i,2}\eta_{i,1}(t)) \text{d} t , \]  
\[ \vdots \]  
\[ \text{d} \eta_{i,n_i + 1}(t) = (\eta_{i,n_i + 1}(t) - k_{i,n_i + 1}\eta_{i,1}(t)) \text{d} t , \]
\[ \text{d} \eta_{i,n_i + 1}(t) = -k_{i,n_i + 1}\eta_{i,1}(t) \text{d} t + \epsilon \Delta_{1,i}(\epsilon t) \text{d} t \]  
\[ + \sqrt{\epsilon} \Delta_{2,i}(\epsilon t) \text{d} \hat{B}_i(t) , \quad i = 1, 2, \ldots, m. \]

By Assumptions 3.1, 3.2, there exists an \( \epsilon \)-independent constant \( C_{2,i} > 0 \) such that
\[ |\Delta_{1,i}(\epsilon t)| + |\Delta_{2,i}(\epsilon t)| \leq C_{2,i} \text{ almost surely for all } t \geq 0 . \]

Consider the positive definite functions \( V_i(\cdot) : \mathbb{R}^{n_i + 1} \to \mathbb{R} \) and \( V(\cdot) : \mathbb{R}^{n + m} \to \mathbb{R} \) given by
\[ V_i(\eta_i) = \eta_i H_i \eta_i^\top, \quad \forall \eta_i \in \mathbb{R}^{n_i + 1}, \quad V(\eta) = \sum_{i=1}^{m} V_i(\eta_i) , \]
\[ \forall \eta = (\eta_1, \ldots, \eta_m) \in \mathbb{R}^{n + m} , \]
where \( H_i \) is specified in (17). By the definition of \( V(\cdot) \) in (32),
\[ \lambda_{\min}(H_i) \|\eta_i\|^2 \leq V_i(\eta_i) \leq \lambda_{\max}(H_i) \|\eta_i\|^2 , \]
\[ \left| \frac{\partial V_i(\eta_i)}{\partial \eta_i} \right| \leq 2 \lambda_{\max}(H_i) \|\eta_i\| , \]
\[ \left| \frac{\partial^2 V_i(\eta_i)}{\partial^{2}\eta_i} \right| \leq 2 \lambda_{\max}(H_i) . \]

By applying Itô's formula to \( V_i(\eta_i(t)) \) with respect to \( t \) along the solution \( \eta_i(t) \) of system (30), we obtain
\[ \text{d} V_i(\eta_i(t)) = \left\{ \sum_{j=1}^{n_i} \frac{\partial V_i(\eta_i(t))}{\partial \eta_{i,j}} (\eta_{i,j+1}(t) - k_{i,j}\eta_{i,1}(t)) \right. \]  
\[ - \frac{\partial V_i(\eta_i(t))}{\partial \eta_{i,n_i + 1}} k_{i,n_i + 1}\eta_{i,1}(t) \right\} \text{d} t \]  
\[ + \epsilon \frac{\partial V_i(\eta_i(t))}{\partial \eta_{i,n_i + 1}} \Delta_{1,i}(\epsilon t) \text{d} t \]  
\[ + \frac{1}{2} \epsilon \frac{\partial^2 V_i(\eta_i(t))}{\partial \eta_{i,n_i + 1}^2} \Delta_{2,i}(\epsilon t) \text{d} t \]  
\[ + \sqrt{\epsilon} \frac{\partial V_i(\eta_i(t))}{\partial \eta_{i,n_i + 1}} \Delta_{2,i}(\epsilon t) \text{d} \hat{B}_i(t) \]  
\[ \triangleq \Delta_{1,i}(t) \text{d} t + \Delta_{2,i}(t) \text{d} B_i(t) , \quad i = 1, 2, \ldots, m. \]

By Young's inequality, we have
\[ 2 \epsilon \lambda_{\max}(H_i) C_{2,i} \|\eta_i(t)\| \leq \frac{1}{2} \|\eta_i(t)\|^2 + 2 \epsilon^2 \lambda_{\max}(H_i) C_{2,i}^2 . \]
This together with (17), (31), (33), and (34) yields that
\[
\frac{d\mathbb{E} V_i(\eta_i(t))}{dt} \leq -\mathbb{E} \| \eta_i(t) \|^2 + 2 \varepsilon \lambda_{\max} (H_i) C_{2,i} \mathbb{E} \| \eta_i(t) \|
+ \varepsilon \lambda_{\max} (H_i)^2 C_{2,i}^2
\]
\[
\leq -\frac{1}{2} \mathbb{E} \| \eta_i(t) \|^2 + 2 \varepsilon^2 \lambda_{\max}^2 (H_i) C_{2,i}^2
+ \varepsilon \lambda_{\max} (H_i)^2 C_{2,i}^2
\]
\[
\leq -\frac{1}{2 \lambda_{\max} (H_i)} \mathbb{E} V_i(\eta_i(t)) + \varepsilon (2 \lambda_{\max}^2 (H_i) C_{2,i}^2
+ \lambda_{\max} (H_i) C_{2,i}^2),
\]
(36)

Set
\[ t_{i,\varepsilon} = 2(n_i + 1) \lambda_{\max} (H_i) \varepsilon \ln \varepsilon^{-1}. \]
(37)

For all \( t \geq t_{i,\varepsilon} \), we have
\[
\mathbb{E} V_i \left( \eta_i \left( \frac{t}{\varepsilon} \right) \right)
\leq e^{-t/(2 \lambda_{\max} (H_i) \varepsilon)} \mathbb{E} V_i(\eta_i(0)) + \varepsilon (2 \lambda_{\max}^2 (H_i) C_{2,i}^2
+ \lambda_{\max} (H_i) C_{2,i}^2) \varepsilon
\]
\[
\times \mathbb{E} \left\| \left( \frac{x_{i,j}(0) - \hat{x}_{i,j}(0)}{\varepsilon^{n_i}}, \ldots, x_{i,n_i+1}(0) - \hat{x}_{i,n_i+1}(0) \right) \right\|^2
+ 2 \lambda_{\max} (H_i) (2 \lambda_{\max}^2 (H_i) C_{2,i}^2 + \lambda_{\max} (H_i) C_{2,i}^2) \varepsilon
\]
\[
\leq \lambda_{\max} (H_i) e^{-t/(2 \lambda_{\max} (H_i) \varepsilon)} \mathbb{E} \left\| \left( \frac{x_{i,j}(0) - \hat{x}_{i,j}(0)}{\varepsilon^{n_i}}, \ldots, x_{i,n_i+1}(0) - \hat{x}_{i,n_i+1}(0) \right) \right\|^2
+ 2 \lambda_{\max} (H_i) (2 \lambda_{\max}^2 (H_i) C_{2,i}^2 + \lambda_{\max} (H_i) C_{2,i}^2) \varepsilon
\]
\[
= \varepsilon \lambda_{\min} (H_i) \Gamma_i,
\]
where
\[
\Gamma_i = \frac{\lambda_{\max} (H_i) \sum_{j=1}^{n_i+1} \mathbb{E} \| x_{i,j}(0) - \hat{x}_{i,j}(0) \|^2}{\lambda_{\min} (H_i)}.
\]
(39)

Therefore, for all \( j = 1, 2, \ldots, n_i + 1 \) and \( t \geq t_{i,\varepsilon} \),
\[
\mathbb{E} \| x_{i,j}(t) - \hat{x}_{i,j}(t) \|^2 \leq e^{2n_i + 2 - 2j} \mathbb{E} \left\| \eta_i \left( \frac{t}{\varepsilon} \right) \right\|^2
\]
\[
\leq e^{2n_i + 2 - 2j} \mathbb{E} \left\| \eta_i \left( \frac{t}{\varepsilon} \right) \right\|^2
\]
\[
\leq e^{2n_i + 2 - 2j} \mathbb{E} V_i(\eta_i \left( \frac{t}{\varepsilon} \right))
\]
\[
\leq \Gamma_i (e^{2n_i + 3 - 2j}),
\]
(40)

where \( \Gamma_i \) is an \( \varepsilon \)-independent constant defined as in (39). This completes the proof of conclusion (i). Next we shall prove the conclusion (ii).

Since \( \mathbb{E} V_i(\eta_i(0)) \) is bounded by \( 1/\varepsilon^{n_i} \) multiplied by an \( \varepsilon \)-independent constant if \( 0 < \varepsilon < 1 \), similar to (36) and (38), for any positive constant \( a > 0 \), we can obtain
\[
\mathbb{E} V_i \left( \eta_i \left( \frac{t}{\varepsilon} \right) \right) \leq e^{-t/(2 \lambda_{\max} (H_i) \varepsilon)} \mathbb{E} V_i(\eta_i(0)) + \varepsilon (2 \lambda_{\max}^2 (H_i) C_{2,i}^2
+ \lambda_{\max} (H_i) C_{2,i}^2)
\]
\[
\times \mathbb{E} \left\| \left( \frac{x_{i,j}(0) - \hat{x}_{i,j}(0)}{\varepsilon^{n_i}}, \ldots, x_{i,n_i+1}(0) - \hat{x}_{i,n_i+1}(0) \right) \right\|^2
+ 2 \lambda_{\max} (H_i) (2 \lambda_{\max}^2 (H_i) C_{2,i}^2 + \lambda_{\max} (H_i) C_{2,i}^2) \varepsilon
\]
\[
\leq e^{-a/(2 \lambda_{\max} (H_i) \varepsilon)} \mathbb{E} V_i(\eta_i(0))
\]
\[
+ 2 \lambda_{\max} (H_i) (2 \lambda_{\max}^2 (H_i) C_{2,i}^2 + \lambda_{\max} (H_i) C_{2,i}^2) \varepsilon
\]
\[
\rightarrow 0 \text{ uniformly in } [a, \infty) \text{ as } \varepsilon \rightarrow 0.
\]
(41)

This concludes that
\[
\mathbb{E} \left| x_{i,j}(t) - \hat{x}_{i,j}(t) \right|^2 \leq \varepsilon^{2n_i + 2 - 2j} \mathbb{E} \left| \eta_i \left( \frac{t}{\varepsilon} \right) \right|^2
\]
\[
\leq \varepsilon^{2n_i + 2 - 2j} \mathbb{E} \left| \eta_i \left( \frac{t}{\varepsilon} \right) \right|^2
\]
\[
\leq \varepsilon^{2n_i + 2 - 2j} \mathbb{E} \left| \eta_i \left( \frac{t}{\varepsilon} \right) \right|^2
\]
\[
\rightarrow 0 \text{ uniformly in } [a, \infty) \text{ as } \varepsilon \rightarrow 0,
\]
\[
\text{for } j = 1, 2, \ldots, n_i + 1,
\]
(42)

which completes the proof of conclusion (ii).

\textbf{Proof of Theorem 3.2:} Let
\[
\eta_{i,j}(t) = \frac{x_{i,j}(t) - \hat{x}_{i,j}(t)}{\varepsilon^{n_i+1-j}}, \quad j = 1, 2, \ldots, n_i + 1,
\]
(43)

and
\[
\eta_i(t) = (\eta_{i,1}(t), \eta_{i,2}(t), \ldots, \eta_{i,n_i+1}(t)).
\]
(44)

A straightforward computation shows that \( \eta_i(t) = (\eta_{i,1}(t), \eta_{i,2}(t), \ldots, \eta_{i,n_i+1}(t)) \) satisfies the following Itô-type stochastic differential equation:
\[
d\eta_{i,1}(t) = (\eta_{i,2}(t) - k_{i,1} (\eta_{i,1}(t))) \, dt,
\]
\[
d\eta_{i,2}(t) = (\eta_{i,3}(t) - k_{i,2} (\eta_{i,1}(t))) \, dt,
\]
\[
\vdots
\]
\[
d\eta_{i,n_i}(t) = (\eta_{i,n_i+1}(t) - k_{i,n_i} (\eta_{i,1}(t))) \, dt,
\]
\[
d\eta_{i,n_i+1}(t) = -k_{i,n_i+1}(t) \eta_{i,1}(t) \, dt + \varepsilon \Delta_{i,1}(t) \, dt
\]
\[
+ \sqrt{\varepsilon} \Delta_{i,2}(t) \, d\phi(t),
\]
(45)

where \( \Delta_{i,1}(t) \) and \( \Delta_{i,2}(t) \) are defined as that in (28). As a consequence of Lemmas 2.3-2.4, the nominal part of system (45)
\[
\hat{\eta}_i(t) = \Phi_i(\eta_i(t))
\]
(46)

is global finite-time stable, where \( \Phi_i(\cdot) \) is defined in (13).

Using Lemma 2.1, we can conclude that there exists a positive definite, radially unbounded, differentiable function \( V_i : \mathbb{R}^{n_i+1} \rightarrow \mathbb{R} \) such that \( V_i(\eta_i) \) is homogeneous of degree \( \theta_i \) with
respect to weights \( \{r_{ij} = (j - 1)a_i - (j - 2)\}_{j=1}^{n_i+1} \), and the Lie derivative of \( V_i \) along the vector fields \( \Phi_j \):

\[
L_{\Phi_j} V_i(\eta_i) = \sum_{j=1}^{n_i} \frac{\partial V_i(\eta_i)}{\partial \eta_{ij}} (\eta_{ij}^{n_i} - k_{ij}[\eta_{ij}^{n_i+1}])
- \frac{\partial^2 V_i(\eta_i)}{\partial \eta_{i,j+1} \partial \eta_{i,j}} k_{i,j+1}[\eta_{i,j+1}^{n_i+1}]^{a_i-1} - \theta_i \tag{47}
\]

is negative definite, where \( \theta_i \) is a constant satisfying \( \theta_i > \max_{1 \leq j \leq n_i+1} (2r_{ij} + d) \) with \( d_i \equiv 1 - a_i \) defined as that in Lemma 2.3.

By the definition of homogeneous of \( V_i \) in Definition 2.2, for any \( \lambda > 0 \), we have

\[
V_i(\lambda^{n_i} \eta_{i1}, \ldots, \lambda^{n_{i,n_i+1}} \eta_{i,n_i+1}) = \lambda^{\theta_i} V_i(\eta_{i1}, \ldots, \eta_{i,n_i+1}). \tag{48}
\]

The derivatives of both sides of (48) with respect to the arguments \( \eta_{ij} \) are given by

\[
\lambda^{\theta_i} \frac{\partial V_i(\lambda^{n_i} \eta_{i1}, \ldots, \lambda^{n_{i,n_i+1}} \eta_{i,n_i+1})}{\partial \eta_{ij}} = \lambda^{\theta_i} \frac{\partial V_i(\eta_{i1}, \ldots, \eta_{i,n_i+1})}{\partial \eta_{ij}}, \quad j = 1, 2, \ldots, n_i + 1. \tag{49}
\]

Thus, \( (\partial V_i(\eta_{i1}, \ldots, \eta_{i,n_i+1}))/\partial \eta_{ij} \) is homogeneous of degree \( \theta_i - r_{ij} \) with respect to weights \( \{r_{ij}\}_{j=1}^{n_i+1} \).

Finding the second partial derivative with respect to \( \eta_{i,n_i+1} \) of both sides of the equation (48) gives

\[
\lambda^{2r_{i,n_i+1}} \frac{\partial^2 V_i(\lambda^{n_i} \eta_{i1}, \ldots, \lambda^{n_{i,n_i+1}} \eta_{i,n_i+1})}{\partial \eta_{i,n_i+1}^2} = \lambda^{\theta_i} \frac{\partial^2 V_i(\eta_{i1}, \ldots, \eta_{i,n_i+1})}{\partial \eta_{i,n_i+1}^2}, \tag{50}
\]

which shows that \( (\partial^2 V_i(\eta_{i1}, \ldots, \eta_{i,n_i+1}))/\partial \eta_{i,n_i+1}^2 \) is homogeneous of degree \( \theta_i - 2r_{i,n_i+1} \) with respect to weights \( \{r_{i,j}\}_{j=1}^{n_i+1} \).

Moreover, the Lie derivative of \( V_i \) along the vector field \( \Phi_i \) satisfies

\[
L_{\Phi_i} V_i(\lambda^{n_i} \eta_{i1}, \ldots, \lambda^{n_{i,n_i+1}} \eta_{i,n_i+1}) = \sum_{j=1}^{n_i+1} \frac{\partial V_i(\lambda^{n_i} \eta_{i1}, \ldots, \lambda^{n_{i,n_i+1}} \eta_{i,n_i+1})}{\partial \eta_{ij}} \Phi_{ij}(\lambda^{n_i} \eta_{i1}, \ldots, \lambda^{n_{i,n_i+1}} \eta_{i,n_i+1})
- \lambda^{\theta_i-d_i} \sum_{j=1}^{n_i+1} \frac{\partial V_i(\eta_{i1}, \ldots, \eta_{i,n_i+1})}{\partial \eta_{ij}} \Phi_{ij}(\eta_{i1}, \ldots, \eta_{i,n_i+1})
- \lambda^{\theta_i-d_i} L_{\Phi_i} V_i(\eta_{i1}, \ldots, \eta_{i,n_i+1}). \tag{51}
\]

Hence \( L_{\Phi_i} V_i \) is homogeneous of degree \( \theta_i - d_i \) with respect to weights \( \{r_{i,j}\}_{j=1}^{n_i+1} \). The above analysis together with Lemma 2.2 yields the following inequalities:

\[
\left| \frac{\partial V_i(\eta_i)}{\partial \eta_{i,n_i+1}} \right| \leq \alpha_i V_i(\eta_i)^{\theta_i - r_{i,n_i+1}/\theta_i}, \quad \forall \eta_i = (\eta_{i1}, \ldots, \eta_{i,n_i+1}) \in \mathbb{R}^{n_i+1}, \tag{52}
\]

\[
\left| \frac{\partial^2 V_i(\eta_i)}{\partial \eta_{i,n_i+1}^2} \right| \leq \beta_i V_i(\eta_i)^{\theta_i - 2r_{i,n_i+1}/\theta_i}, \quad \forall \eta_i = (\eta_{i1}, \ldots, \eta_{i,n_i+1}) \in \mathbb{R}^{n_i+1}, \tag{53}
\]

and

\[
L_{\Phi_i} V_i(\eta_i) \leq -\gamma_i V_i(\eta_i)^{\theta_i - d_i/\theta_i}, \quad \forall \eta_i = (\eta_{i1}, \ldots, \eta_{i,n_i+1}) \in \mathbb{R}^{n_i+1}, \tag{54}
\]

where \( \alpha_i, \beta_i, \gamma_i \) are some positive constants.

From Assumptions 3.1, 3.2, similar to (31), there exists an \( \varepsilon \)-independent constant \( C_2i > 0 \) such that

\[
|\Delta_{1,i}(\varepsilon t)| + |\Delta_{2,i}(\varepsilon t)| \leq C_2i \tag{55}
\]

Applying Itô’s formula to \( V_i(\eta_i(t)) \) with respect to \( t \) along the solution \( \eta_i(t) \) of system (45) gives

\[
\frac{dV_i(\eta_i(t))}{dt} = L_{\Phi_i} V_i(\eta_i(t)) + \frac{\partial V_i(\eta_i(t))}{\partial \eta_{i,n_i+1}} \Delta_{1,i}(\varepsilon t) dt
+ \frac{\varepsilon}{2} \left( \frac{\partial^2 V_i(\eta_i(t))}{\partial \eta_{i,n_i+1}^2} \Delta_{2,i}(\varepsilon t) dt
+ \sqrt{\varepsilon} \frac{\partial V_i(\eta_i(t))}{\partial \eta_{i,n_i+1}} \Delta_{2,i}(\varepsilon t) d\hat{B}_i(t) \right), \tag{56}
\]

where \( L_{\Phi_i} V_i(\eta_i(t)) \) is given in (47). It then follows from (52), (53), (54), (55), and (56) that

\[
\frac{d\mathbb{E}V_i(\eta_i(t))}{dt} \leq \mathbb{E}L_{\Phi_i} V_i(\eta_i(t)) + \varepsilon \mathbb{E} \left| \frac{\partial V_i(\eta_i(t))}{\partial \eta_{i,n_i+1}} \Delta_{1,i}(\varepsilon t) \right|
+ \frac{\varepsilon}{2} \mathbb{E} \left| \frac{\partial^2 V_i(\eta_i(t))}{\partial \eta_{i,n_i+1}^2} \Delta_{2,i}(\varepsilon t) \right|
\leq -\gamma_i \mathbb{E} V_i(\eta_i(t))^{\theta_i-d_i/\theta_i}
+ \varepsilon \alpha_i C_2i \varepsilon \mathbb{E} V_i(\eta_i(t))^{\theta_i - r_{i,n_i+1}/\theta_i}
+ \frac{\varepsilon \beta_i C_2i}{2} \mathbb{E} V_i(\eta_i(t))^{\theta_i - 2r_{i,n_i+1}/\theta_i}. \tag{57}
\]

Choose a positive constant \( \mu_i \) satisfying \( \mu_i < (\gamma_i(\theta_i - d_i))/((2\theta_i - 3r_{i,n_i+1}))/2 > 0 \). Then

\[
\xi_i \equiv \gamma_i - \frac{2\theta_i - 3r_{i,n_i+1}}{\theta_i - d_i} > 0. \tag{58}
\]

Since \( a_i > 1 - (1/(n_i + 1)) \), a simple computation shows that \( r_{i,n_i+1} > d_i \). Thus \( \theta_i > r_{i,n_i+1} > d_i \) and \( (\theta_i - d_i)/(\theta_i - r_{i,n_i+1}) \)
> 1, \((\theta_i - d_i)/(r_{i,n+1} - d_i)\) > 1. By Young's inequality, it is easy to check that
\[
\varepsilon \alpha C_2 i (V_i(\eta_i))^{(\theta_i - r_{i,n+1})/\theta_i} \leq \mu_i \frac{\theta_i - r_{i,n+1}}{\theta_i - d_i} (V_i(\eta_i))^{(\theta_i - d_i)/\theta_i} \\
+ \frac{r_{i,n+1} - d_i}{\theta_i - d_i} \mu_i \left( \frac{1}{r_{i,n+1} - r_{i,n+1} - d_i} \right) \\
\times (\varepsilon \alpha C_2 i (V_i(\eta_i))^{(\theta_i - d_i)/(r_{i,n+1} - d_i)}).
\] (59)

On the other hand, since \(\theta_i > 2r_{i,n+1} > d_i\), we have \((\theta_i - d_i)/(\theta_i - 2r_{i,n+1} - d_i) > 1\), \((\theta_i - 2r_{i,n+1} - d_i)/(2r_{i,n+1} - d_i) > 1\). Again, it follows from Young's inequality that
\[
\frac{\varepsilon \beta C_2 i}{2} (V_i(\eta_i))^{(\theta_i - 2r_{i,n+1})/\theta_i} \leq \mu_i \frac{\theta_i - 2r_{i,n+1}}{\theta_i - d_i} (V_i(\eta_i))^{(\theta_i - d_i)/\theta_i} \\
+ \frac{2r_{i,n+1} - d_i}{\theta_i - d_i} \mu_i \left( \frac{1}{2r_{i,n+1} - 2r_{i,n+1} - d_i} \right) \\
\times \left( \frac{\varepsilon \beta C_2 i}{2} \right)^{(\theta_i - d_i)/(2r_{i,n+1} - d_i)}.
\] (60)

Set
\[
q_i = \frac{r_{i,n+1} - d_i}{\theta_i - d_i} \mu_i \left( \frac{1}{r_{i,n+1} - r_{i,n+1} - d_i} \right) (\alpha_i C_2 i) (\theta_i - d_i)/(r_{i,n+1} - d_i) \\
+ \frac{2r_{i,n+1} - d_i}{\theta_i - d_i} \mu_i \left( \frac{1}{2r_{i,n+1} - 2r_{i,n+1} - d_i} \right) \\
\times \left( \frac{\varepsilon \beta C_2 i}{2} \right)^{(\theta_i - d_i)/(2r_{i,n+1} - d_i)}.
\] (61)

Suppose that \(0 < \varepsilon < 1\). Considering (57)–(61). We then obtain
\[
\frac{d\mathbb{E}[V_i(\eta_i(t))]}{dt} \leq -\xi_i \mathbb{E}[V_i(\eta_i(t))]^{(\theta_i - d_i)/\theta_i} \\
+ q_i \mathbb{E}^{(\theta_i - d_i)/(2r_{i,n+1} - d_i)}.
\] (62)

Now, we choose \(\Lambda_i\) satisfying
\[
1 < \Lambda_i < \frac{\theta_i - d_i}{2r_{i,n+1} - d_i}.
\] (63)

Suppose that \(0 < \varepsilon < \varepsilon^* \triangleq \min\{1, \xi_i/\theta_i (\theta_i - d_i - \Lambda_i)(2r_{i,n+1} - d_i)\}\). Therefore,
\[
\sigma_{i,e} \triangleq \xi_i - q_i \mathbb{E}^{(\theta_i - d_i)/(2r_{i,n+1} - d_i - \Lambda_i)} > 0.
\] (64)

If \(\mathbb{E}[V_i(\eta_i(t))]^{(\theta_i - d_i)/\theta_i} \geq e^{\Lambda_i t}\), from (62), it follows that
\[
\frac{d\mathbb{E}[V_i(\eta_i(t))]}{dt} \leq -e^{\Lambda_i} \sigma_{i,e} < 0.
\] (65)

According to Hölder's inequality, we can easily verify that
\[
\mathbb{E}(V_i(\eta_i(t)))^{(\theta_i - d_i)/\theta_i} \leq (\mathbb{E}(V_i(\eta_i(t)))^{(\theta_i - d_i)/\theta_i})^{(\theta_i - d_i)/\theta_i}.
\] (66)

This together with (65) yields that there exist an \(\varepsilon\)-dependent constant \(t_{i,e}\) such that for all \(t \geq t_{i,e}\),
\[
\mathbb{E}(V_i(\eta_i(t)))^{(\theta_i - d_i)/\theta_i} \leq e^{\Lambda_i}.
\] (67)

We notice that \(|\eta_{i,j}|^2\) as the function of \((\eta_{i,1}, \ldots, \eta_{i,j}, \ldots, \eta_{i,n+1})\) is a homogeneous function of degree \(2r_{i,j}\) with respect to weights \(\{r_{i,j}\}_{j=1}^{n+1}\). It follows from Lemma 2.2 that there exists a positive constant \(\Gamma_{i,j}\) such that
\[
|\eta_{i,j}|^2 \leq \Gamma_{i,j}(V_i(\eta_i))^{(2r_{i,j})/\theta_i}.
\] (68)

It is easy to check that \((\theta_i - d_i)/(2r_{i,n+1} - d_i) \cdot (2r_{i,j}/(\theta_i - d_i)) = (2r_{i,j})/(2r_{i,n+1} - d_i) > 1\). This together with (63) yields that there exists \(\Lambda_i\) such that
\[
\Lambda_{i,j} \triangleq \Lambda_i \cdot \frac{2r_{i,j}}{\theta_i - d_i} > 1.
\] (69)

Since \(\theta_i > 2r_{i,j} + d_i\), we have \(2r_{i,j}/(\theta_i - d_i) < 1\). Using Hölder's inequality, we can obtain that
\[
\mathbb{E}[V_i(\eta_i(t))]^{2r_{i,j}/\theta_i} \leq (\mathbb{E}[V_i(\eta_i(t))]^{(\theta_i - d_i)/\theta_i})^{2r_{i,j}/(\theta_i - d_i)}.
\] (70)

Thus, it follows from (67)–(70) that there exists an \(\varepsilon\)-dependent constant \(t_{i,e}\) such that for all \(t \geq t_{i,e}\),
\[
\mathbb{E}[x_{i,j}(t) - \hat{x}_{i,j}(t)]^2 \leq e^{2\varepsilon^{n+2+2\varepsilon^*}(\eta_{i,j})^2} \\
\leq \Gamma_{i,j} e^{2\varepsilon^{n+2+2\varepsilon^*}(\mathbb{E}[V_i(\eta_i(t))])}^{2r_{i,j}/\theta_i} \\
\leq \Gamma_{i,j} e^{2\varepsilon^{n+2+2\varepsilon^*}(\eta_{i,j})^2},
\] (71)

where \(\Gamma_{i,j}\) is an \(\varepsilon\)-independent constant and \(\Lambda_{i,j} > 1\) that are specified in (68) and (69), respectively. This completes the proof of Theorem 3.2.

\[\square\]

5. Numerical simulations

In this section, we present an example to illustrate the effectiveness of the proposed linear ESO and homogeneous ESO numerically. Consider the following uncertain MIMO system with stochastic uncertainties:
\[
\dot{x}_{1,1}(t) = x_{1,2}(t),
\]
\[
x_{1,2}(t) = c_1 x_{1,2}(t) + c_2 \cos(x_{2,1}^2(t) + x_{2,2}^2(t)) \\
+ w_1(t) + u_1(t),
\]
\[
\dot{x}_{2,1}(t) = x_{2,2}(t),
\]
\[
\dot{x}_{2,2}(t) = c_3 x_{2,2}(t) + c_4 \sin(x_{1,2}^3(t)) + w_2(t) + u_2(t),
\]
\[
y_1(t) = x_{1,1}(t),
\]
\[
y_2(t) = x_{2,1}(t),
\]

where \(c_i (i = 1, 2, \ldots, 8)\) are unknown parameters satisfying: \(|c| \leq M\) for any given (known) constant \(M > 0\), and \(c_1 < 0, c_3 < 0\). The \(w_1(t) \triangleq \cos(c_2 t + c_6 B_1(t))\) and \(w_2(t) \triangleq \cos(c_2 t + c_8 B_2(t))\) are bounded non-white noise appeared often in many practical dynamical systems (Hu et al., 2012; Huang & Zhu, 2004; Huang et al., 2002; Li et al., 2008), where \(c_5, c_7\) are constants representing the central frequency, and \(c_6, c_8\) are strengths of frequency disturbance. In this case, \(k = m = n_1 = n_2 = 2\).

The solution of system (72) may not be bounded almost surely. However, we apply direct output feedbacks \(u_1(t) = -y_1(t) + v_1(t)\) and \(u_2(t) = -y_2(t) + v_2(t)\), where \(v_1(t)\) and \(v_2(t)\) are the new control inputs, and for simplicity, we just take
$v_1(t) = v_2(t) = 0$. In this case, it is easy to check that Assumptions 3.1 and 3.2 are satisfied. It then follows from Theorem 3.1 that we can design a linear ESO (73) for system (72) as follows:

$$\dot{x}_{1,1}(t) = \dot{x}_{1,2}(t) + \frac{3}{\varepsilon} (y_1(t) - \hat{x}_{1,1}(t)),$$

$$\dot{x}_{1,2}(t) = \dot{x}_{1,3}(t) + \frac{3}{\varepsilon} (y_1(t) - \hat{x}_{1,1}(t)) + u_1(t),$$

$$\dot{x}_{1,3}(t) = \frac{1}{\varepsilon} \gamma_a \left[ \frac{y_1(t) - \hat{x}_{1,1}(t)}{\varepsilon^2} \right]^{3a-2} + u_1(t),$$

$$\dot{x}_{2,1}(t) = \dot{x}_{2,2}(t) + \frac{3}{\varepsilon} (y_2(t) - \hat{x}_{2,1}(t)),$$

$$\dot{x}_{2,2}(t) = \dot{x}_{2,3}(t) + \frac{3}{\varepsilon} (y_2(t) - \hat{x}_{2,1}(t)) + u_2(t),$$

$$\dot{x}_{2,3}(t) = \frac{1}{\varepsilon} \gamma_a \left[ \frac{y_2(t) - \hat{x}_{2,1}(t)}{\varepsilon^2} \right]^{3a-2} + u_2(t).$$

Similarly, by Theorem 3.2, we can design a homogeneous ESO as follows:

$$\dot{x}_{1,1}(t) = \dot{x}_{1,2}(t) + \frac{3}{\varepsilon} (y_1(t) - \hat{x}_{1,1}(t))^a,$$

$$\dot{x}_{1,2}(t) = \dot{x}_{1,3}(t) + \frac{3}{\varepsilon} (y_1(t) - \hat{x}_{1,1}(t))^{a-1} + u_1(t),$$

$$\dot{x}_{1,3}(t) = \frac{1}{\varepsilon} \left[ \frac{y_1(t) - \hat{x}_{1,1}(t)}{\varepsilon^2} \right]^{3a-2},$$

$$\dot{x}_{2,1}(t) = \dot{x}_{2,2}(t) + \frac{3}{\varepsilon} (y_2(t) - \hat{x}_{2,1}(t))^a,$$

$$\dot{x}_{2,2}(t) = \dot{x}_{2,3}(t) + \frac{3}{\varepsilon} (y_2(t) - \hat{x}_{2,1}(t))^{a-1} + u_2(t),$$

$$\dot{x}_{2,3}(t) = \frac{1}{\varepsilon} \left[ \frac{y_2(t) - \hat{x}_{2,1}(t)}{\varepsilon^2} \right]^{3a-2}.$$

First we notice that the corresponding matrices in (16) for both (73) and (74)

$$E_1 = E_2 = \begin{pmatrix} -3 & 1 & 0 \\ -3 & 0 & 1 \\ -1 & 0 & 0 \end{pmatrix} \quad (75)$$

for which all eigenvalues equal to -1 are Hurwitz.

The stochastic total disturbances in two subsystems are given by

$$x_{1,3}(t) = c_1 x_{1,2}(t) + c_2 \cos(x_{1,1}^2(t) + x_{2,2}(t)) + w_1(t),$$

$$x_{2,3}(t) = c_3 x_{2,2}(t) + c_4 \sin(x_{1,2}^2(t)) + w_2(t). \quad (76)$$

Figures 2–5 display the numerical results for (72)–(74) where we take the initial values as

$$x_{1,1}(0) = x_{2,1}(0) = 1, \quad x_{1,2}(0) = x_{2,2}(0) = -1,$$

$$\dot{x}_{1,1}(0) = \dot{x}_{1,2}(0) = \dot{x}_{2,1}(0) = \dot{x}_{2,2}(0) = 0, \quad (77)$$

the uncertain parameters as

$$c_1 = -2, \quad c_2 = 4, \quad c_3 = c_4 = 2, \quad c_5 = c_7 = 1,$$

$$c_6 = 2, \quad c_8 = 3, \quad (78)$$

time discrete step $\Delta t$ as $\Delta t = 0.001$, the tuning parameter $\varepsilon$ as $\varepsilon = 0.05$, and the parameter $a$ in homogeneous ESO (74) as $a = 0.8$.

It is observed from Figures 2 and 5 that the linear ESO (73) and homogeneous ESO (74) are very effective in estimating the states $(x_{1,1}(t), x_{1,2}(t), x_{2,1}(t), x_{2,2}(t))$ and the stochastic total disturbance $(x_{1,3}(t), x_{2,3}(t))$ defined as that in (76), respectively. Theoretically, we can conclude from Theorem 3.1 that the estimation errors of the linear ESO (73) for $x_{i,1}(t)$ ($i = 1, 2$), $x_{i,2}(t)$ ($i = 1, 2$), $x_{i,3}(t)$ ($i = 1, 2$) are bounded by $O(\varepsilon^5)$, $O(\varepsilon^5)$, $O(\varepsilon)$ in practical mean square sense, respectively. It is observed

Figure 2. Estimation of states $(x_{1,1}(t), x_{1,2}(t), x_{2,1}(t), x_{2,2}(t))$ and stochastic total disturbance $(x_{1,3}(t), x_{2,3}(t))$ by linear ESO (73).
from Figure 2 that the estimation effect for $x_{i,1}(t)$ ($i = 1, 2$) is the best, $x_{i,2}(t)$ ($i = 1, 2$) the second, and $x_{i,3}(t)$ ($i = 1, 2$) the last, which are coincident with the theoretical estimations. In addition, we can conclude from Theorem 3.2 that the estimation errors of the homogeneous ESO (74) for $x_{i,1}(t)$ ($i = 1, 2$), $x_{i,2}(t)$ ($i = 1, 2$), $x_{i,3}(t)$ ($i = 1, 2$) are bounded by $O(e^{\Delta_{4}t})$, $O(e^{\Delta_{3}t})$, $O(e^{\Delta_{2}t})$ in practical mean square sense, respectively, where $\Delta_{j} > 1$ ($j = 1, 2, 3$). It is seen from Figure 4 that the estimation effect for $x_{i,1}(t)$ ($i = 1, 2$) is the best, $x_{i,2}(t)$ ($i = 1, 2$) the second, and $x_{i,3}(t)$ ($i = 1, 2$) the last, which are coincident with the theoretical estimations.

The main problem for high gain ESO, likewise many other high gain designs, is the peaking value problem near the initial stage caused by different initial values of system (72) and ESO. It is observed from Figure 3 that the absolute values of states $\hat{x}_{1,2}(t)$, $\hat{x}_{1,1}(t)$, $\hat{x}_{2,2}(t)$, and $\hat{x}_{2,3}(t)$ of linear ESO are near 100, 3000, 100, and 3000 near the initial stage, respectively. However, it is seen from Figure 5 that the absolute values of states $\hat{x}_{1,2}(t)$, $\hat{x}_{1,1}(t)$, $\hat{x}_{2,2}(t)$, and $\hat{x}_{2,3}(t)$ of homogeneous ESO near the initial stage have much smaller peaking values than those of linear ESO being only 20, 100, 20, and 100, respectively.

Finally, we see that the homogeneous ESO can estimate the stochastic total disturbances $(x_{3,1}(t), x_{3,2}(t))$ with smaller errors than linear ESO by comparing Figures 2(c) and 2(f) with 4(c) and 4(f), which validates the Remark 3.3 numerically.
6. Concluding remarks and future work

In this paper, both linear ESO and homogeneous ESO are designed for a class of MIMO nonlinear systems composed of coupled subsystems subject to large stochastic uncertainties. These two classes of ESOs are used to estimate not only the state but also the stochastic total disturbance of each subsystem including internal coupled unmodelled dynamics and external stochastic disturbance without known statistical characteristics. We give rigorous theoretical proofs that the estimation errors ultimately converge to an arbitrarily small neighbourhood of zero in mean square sense with the estimation performance guaranteed by suitably choosing the high gain parameter of ESO. In addition, the theoretical results show that homogeneous ESO can achieve higher estimation accuracy than linear ESO. Numerical simulations show that the ESOs are very satisfactory in estimating state and stochastic total disturbance of each subsystem, and the nonlinear ESO has smaller peaking values and more accurate estimation than the linear ESO.

Finally, we indicate that a future work would be the performance analysis of the resulting closed-loop systems under ESO-based feedback controls and the applications of ESO or ADRC approach to more general stochastic nonlinear systems like system (7) driven by the Brownian motions.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was carried out with the support of the National Natural Science Foundation of China [grant numbers 11501108, 11501106], the Project of Department of Education of Guangdong Province [grant numbers 2017KTSCX191, 2017KZDXM087, 2014KZDXM063], and the Natural Science Foundation of Guangdong Province [grant numbers 2015A030313636, 2016A030313835].

ORCID

Ze-Hao Wu http://orcid.org/0000-0003-4109-1757
Bao-Zhu Guo http://orcid.org/0000-0001-9078-0001

References


