Asymptotic and tracking guarantees in interval observer design for systems with unmeasured polytopic nonlinearities
Hiroshi Ito, Thach Ngoc Dinh

To cite this version:
Hiroshi Ito, Thach Ngoc Dinh. Asymptotic and tracking guarantees in interval observer design for systems with unmeasured polytopic nonlinearities. 21st IFAC World Congress, Jul 2020, Berlin, Germany. hal-02516113

HAL Id: hal-02516113
https://hal-cnam.archives-ouvertes.fr/hal-02516113
Submitted on 23 Mar 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Asymptotic and tracking guarantees in interval observer design for systems with unmeasured polytopic nonlinearities*

Hiroshi Ito* Thach Ngoc Dinh**

* Department of Intelligent and Control Systems, Kyushu Institute of Technology, Fukuoka 820-8502, Japan (e-mail: hiroshi@ces.kyutech.ac.jp)
** CNAM, Cedric-Lab, 292 rue St-Martin, 75141 Paris Cedex 03, France (e-mail: ngoc-thach.dinh@lecnam.net)

Abstract: For a class of nonlinear systems subject to disturbances, an observer is proposed to estimate time-varying intervals in which their state variables are guaranteed to stay all the time. The objective is to effectively deal with nonlinearities depending on unmeasured variables which have been usually treated as uncertainty in observer design. Focusing on nonlinearities in a polytopic form, this paper shows how an interval observer can replace nonlinearities in unmeasured variables by nonlinearities in estimated intervals. Theoretical guarantees and simulation comparisons are presented to demonstrate that the use of interval-dependent nonlinearities gives better estimates than the use of an overbounding observer.

Keywords: | Interval observers; Polytopic systems; State estimation; Nonlinear systems.

1. INTRODUCTION

Recently, interval observers have gained an increased attention as a means of estimating variables component-wise all the time in the presence of disturbances (Gonzé et al. (2000); Bernard and Gonzé (2004); Moisan et al. (2009)). The interval estimation overcomes weak points of classical observers which are only able to give asymptotic estimates in the absence of disturbances. Naturally, such interval observer design is successful in doing this at the cost of restrictive assumptions. Indeed, the interval property requires error systems to be positive, which reduces to Metzler matrices in the case of linear systems. Although some relaxing techniques are available, securing the positivity at some point during the design process remains the key. This paper does not look for better tricks in achieving the positivity. Instead, this paper seeks benefits of interval estimation in dealing with nonlinearities. It is typical of observer design to cancel nonlinearities in the plant for obtaining tractable error dynamics. It amounts to requiring the nonlinearities to be in only measured variables (Krener and Isidori (1983); Isidori (2001); Khalil (2015); Nijmeijer and van der Schaft (1990); Bernard (2019)). The same idea applies to interval observers (Raissi et al. (2012); Elimov et al. (2013); Zheng et al. (2016)). Components involving unmeasured variables are treated as uncertainties there.

Unlike classical observers, interval observers give valid estimates of state variables at each time. The central idea of this paper is to exploit the estimated intervals in observer gains for allowing unmeasured variables to be involved in nonlinearities. To realize this idea, this paper focuses on a polytopic formulation which interpolates linear systems with scalar-valued nonlinear functions of state variables. The formulation is considered as a class of linear parameter-varying (LPV) systems (e.g., Apkarian et al. (1995)), but the parameters are endogenous, i.e., the state variables, which means that the systems are nonlinear. Thus the feedback induced by the endogenous variables should be treated differently from time-varying cases. In particular, in observer design, the endogenous variables, i.e., the state variables, are not measured. The polytopic formulation of nonlinearities is popular in control engineering (Takagi and Sugeno (1985)), which is often called the Takagi-Sugeno fuzzy (T-S) model. Each linear system in the polytopic model can be regarded as a linear model at an operating point.

There have been extensive studies on interval observers in various settings of systems (e.g., Mazenc and Bernard (2011); Mazenc and Dinh (2014); Mazenc et al. (2013, 2014) to name a few in addition to the aforementioned references). Some researches have also been focused on LPV and T-S type systems (Chebotarev et al. (2015); Elimov and Raissi (2016); Li et al. (2019)). In Marx et al. (2019), for discrete-time polytopic model, nonlinearities of unmeasured variables are merely treated as one of external disturbances with respect to which $L_2$ gain is reduced. It is worth noting that the approaches proposed in Raissi et al. (2012) and Elimov et al. (2013) incorporate nonlinearities of unmeasured variables into uncertainties bounded by functions of the input and estimated intervals of state variables. However, the uncertainty treatment does not allow the estimated nonlinearities to be used in the observer gain to make use of the polytopic model. The objective of this paper is to propose a mechanism making use of intervals in the observer to treat nonlinearities of unmeasured variables, and to characterize guarantees one can gain from that mechanism.

Notation The sets of real numbers and non-negative real numbers are denoted by $\mathbb{R}$ and $\mathbb{R}_+$, respectively. The symbol $\| \cdot \|$ denotes Euclidean norm of vectors of any dimension. A square matrix $A \in \mathbb{R}^{n \times n}$ is said to be positive definite and written as $A > 0$ if $v^T Av > 0$.
consider the system whose state \( x(t) \in \mathbb{R}^n \) is governed by

\[
\dot{x}(t) = \sum_{k=1}^{m} \mu_k(x(t), u(t)) A_k x(t) + \beta(y(t), u(t)) + \delta(t)
\]

\[
y(t) = C x(t),
\]

where \( u(t) \in \mathbb{R}^p \) and \( y(t) \in \mathbb{R}^q \) are the input and the output, respectively. This paper refers to (1) as the plant. Assume that \( \mu_k : \mathbb{R}^n \times \mathbb{R}^p \to [0, 1] \) satisfies

\[
\sum_{k=1}^{m} \mu_k(x, u) = 1, \quad \forall x \in \mathbb{R}^n.
\]

Let \( x_i \) denote the \( i \)-th component of the vector \( x \). The objective of this paper is to estimate \( x(t) \) which is not measured. The output \( y(t) \) is measured. We are interested in not only asymptotic estimation of \( x(t) \), but also estimation of time-varying intervals in which all the variables \( x_i(t) \) stay all times. For this purpose, the functions \( \beta : \mathbb{R}^q \times \mathbb{R}^p \to \mathbb{R}^n \) and \( \mu_k : \mathbb{R}^n \times \mathbb{R}^p \to [0, 1] \) are assumed to be locally Lipschitz. For an arbitrary disturbance \( \delta : \mathbb{R}^n \to \mathbb{R}^n \) which is Lebesgue measurable locally essentially bounded, it is assumed that system (1a) is forward complete \(^1\) in order to write properties of state estimation for the infinite time horizon. The forward complete assumption is not needed if one stops evaluating properties at the finite escape time of maximal solutions. Assume that we know the piecewise continuous functions \( \delta, \tilde{\delta} : \mathbb{R}^n \to \mathbb{R}^n \) satisfying

\[
\delta(t) \leq \tilde{\delta}(t), \quad \text{a.e. } t \in \mathbb{R}^+.
\]

The system (1a) with initial conditions restricted to \( x(0) \in \mathbb{R}^n_+ \) is said to be non-negative if there is no instant \( t \in \mathbb{R}^+ \) such that \( x(t) \not\geq 0 \). Note that whenever the system (1a) is said to be non-negative, the restriction \( x(0) \in \mathbb{R}^n_+ \) and \( \delta(t) \geq 0 \) for all \( t \in \mathbb{R}^+ \) are assumed in this paper. Let \( [A_k] \) denote the \( i \)-th row of the matrix \( A_k \). The non-negativity holds if and only if for each \( i \in \{1, 2, \ldots, n\} \), the implication

\[
x_i = 0 \Rightarrow \forall u \in \mathbb{R}^p \sum_{k=1}^{m} \mu_k(x, u) [A_k]_i x + \beta_i(C x, u) \geq 0
\]

holds true for all \( x \in \mathbb{R}^n_+ \). Systems are often non-negative. Indeed, systems are non-negative if their state variables are chosen as mass, i.e., energy quantities.

\(^1\) If \( \beta \) is globally Lipschitz, system (1a) is guaranteed to be forward complete since (2) implies global Lipschitzness of \( \mu_k(s) \).

This paper proposes the following pair of systems as an interval observer:

\[
\hat{x}(t) = \sum_{k=1}^{m} \left\{ \overline{w}_k(z(t), \overline{x}(t), y(t), u(t)) A_k \overline{x}(t) \right\} + \beta(y(t), u(t)) + \tilde{\delta}(t)
\]

\[
\tilde{x}(t) = \sum_{k=1}^{m} \left\{ \underline{w}_k(z(t), \overline{x}(t), y(t), u(t)) A_k \overline{x}(t) \right\} + \beta(y(t), u(t)) + \tilde{\delta}(t),
\]

where the matrices \( L_k, \overline{u}_k \in \mathbb{R}^{n \times q} \) (5) have yet to be determined. The functions \( \overline{w}_k \) and \( \underline{w}_k \) are any bounded continuous functions satisfying

\[
\sum_{k=1}^{m} \mu_k(s, u) A_k \leq \sum_{k=1}^{m} \overline{w}_k(z, \tau, C s, u) A_k, \quad \forall s \in [\min \{z, \tau\}, \tau]
\]

\[
\sum_{k=1}^{m} \underline{w}_k(z, \tau, C s, u) A_k \leq \sum_{k=1}^{m} \mu_k(s, u) A_k, \quad \forall s \in [\max \{z, \tau\}, \tau]
\]

for all \( z, \tau, x \in \mathbb{R}^n, y \in \mathbb{R}^q \) and \( u \in \mathbb{R}^p \). When (1a) is a non-negative system, the domains of \( s \) in (6a) and (6b) are \( [\min \{\max \{0, z\}, \tau\}, \tau] \) and \( [\max \{0, z\}, \max \{\min \{0, z\}, \tau\}] \), respectively. The same applies to the rest of this paper when the non-negativity is assumed for (1a).

Both equations (5a) and (5b) are natural extension of the Luenberger-type observer (see, e.g., Krener and Isidori (1983); Nijmeijer and van der Schaft (1990)), except that the system coefficients depend on \( \tau \) and \( z \). It is stressed that the standard Luenberger-type observer cannot be applied directly to the plant (1a) since the nonlinearities in (1a) involve the unmeasured state \( x \). To deal with such nonlinearities, \( \overline{w}_k \) and \( \underline{w}_k \) are introduced to (5).

Remark 1. In view of modeling a system, the unity restriction (2) for (1) is equivalent to assuming boundedness of \( \mu_k \). Indeed, if functions \( \mu_k : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}_+ \) admit the existence of \( c \geq 0 \) such that \( \mu_k(x, u) \leq c \) holds for all \( x \in \mathbb{R}^n_+ \) and \( u \in \mathbb{R}^p \), the non-zero bound \( c \) can be absorbed by \( A_k \), and the identity of (2) can be achieved by introducing a zero matrix to the set of \( A_k \)s for \( \mu_k(x, u) \) playing the complementary role. When \( \beta(y, u) \) is restricted to \( \beta(y, u) = \sum_{k=1}^{m} \mu_k(x, u) B_k u \), the model (1a) defined with (2) is sometimes referred to as the polytopic linear model in the framework of polytopic linear parameter-varying systems (Apkarian et al. (1995)). The model is also called the Takagi-Sugeno Fuzzy model (Takagi and Sugeno (1985)) which is very popular for engineering nonlinearities in controller design problems. This paper supposes that \( \beta(y, u) \) depends on \( y \) instead of \( x \) in the polytopic formulation. It means that \( B_k \)s are assumed to delete the unmeasurable part from \( \mu_k(x, u) \).

Remark 2. For any plant given by (1a) with (2), the existence of bounded continuous functions \( \overline{w}_k \) and \( \underline{w}_k \) satisfying (6) is guaranteed. More precisely, if achieving (6a)-(6c) directly is not obvious, one can always increase the number \( m \) in (1a) without changing the system dynamics to find...
The matrices $\overline{w}_k$ and $\underline{w}_k$. The easiest approach is to divide the $k$-th mode into two modes (denoted by $k_+, k_-$ below) as
\[
\mu_k(x, u)A_k = \hat{\mu}_k(x, u)A_k + \overline{\mu}_k(x, u)A_k,
\]
\[
\hat{\mu}_k(x, u) = \underline{\mu}_k(x, u) = \frac{1}{2}\mu_k(x, u)
\]
\[
A_k = 2\max\{A_k, 0\}, \quad A_k = -2A_k - A_k.
\]

It is easy to see that
\[
\overline{w}_k^+(\underline{w}, \overline{w}, u)A_k + \overline{w}_k^-(\underline{w}, \overline{w}, u)A_k - \overline{\mu}_k(s, u)A_k + \overline{\mu}_k(s, u)A_k - \overline{w}_k^+(\underline{w}, \overline{w}, u)A_k + \overline{w}_k^-(\underline{w}, \overline{w}, u)A_k - \overline{w}_k^+(\underline{w}, \overline{w}, u)A_k + \overline{w}_k^-(\underline{w}, \overline{w}, u)A_k
\]
hold for $s$ in the ranges of (6a) and (6b), respectively, with
\[
\overline{w}_k^+(\underline{w}, \overline{w}, u) = \overline{w}_k^-(\underline{w}, \overline{w}, u) = \frac{1}{2} s \in [\min(\underline{w}, \overline{w})], \\mu_k(s, u)
\]

These choices meet (6c) and (6d). The boundedness of $\overline{w}_k$ and $\underline{w}_k$ is clear from (2).

3. MAIN RESULT

The goal of this paper is to show $\overline{v}(t)$ and $\underline{v}(t)$ generated by (5) give an upper bound and a lower bound for all the time $t$, respectively, and to provide useful properties suggesting the advantage of using (5). To this end, let $\overline{v} = \overline{v} - x$ and $\underline{v} = \underline{v} - x$, the following is the main result.

**Theorem 3.** Consider the system consisting of (1) and (5) equipped with (2) and (6c). Suppose that there exist $\overline{v}, \underline{v} > 0, \overline{\xi}, \underline{\xi} \geq 0$, $\overline{P}, \underline{P} \geq 0$ and $\overline{P}, \underline{P} \geq 0$ satisfying
\[
\begin{align*}
\sum_{k=1}^{m} \overline{\xi}_k \overline{w}_k(x, \overline{w}, y, u) & \geq \overline{v}, \\
\sum_{k=1}^{m} \underline{\xi}_k \underline{w}_k(x, \overline{w}, y, u) & \geq \underline{v}, \\
\overline{P}(A_k + 2L_k C) + (A_k + 2L_k C)^T \overline{P} + \overline{\xi}_k \overline{P} \leq 0 \\
\underline{P}(A_k + 2L_k C) + (A_k + 2L_k C)^T \underline{P} + \underline{\xi}_k \underline{P} \leq 0
\end{align*}
\]
for all $k = 1, 2, ..., m$. Then the implication
\[
\begin{align*}
\lim_{t \to \infty} \overline{v}(t) - \underline{v}(t) & = \lim_{t \to \infty} x(t) = 0
\end{align*}
\]
holds true of all $x(0), \overline{v}(0), \underline{v}(0) \in \mathbb{R}^n$, and there exist $\overline{v}, \underline{v} \geq 0$ such that
\[
\begin{align*}
\limsup_{t \to \infty} |\overline{v}(t) - \underline{v}(t)| & = (\overline{\sigma} + \underline{\sigma}) \limsup_{t \to \infty} |x(t)| \\
& + \overline{\sigma} \limsup_{t \to \infty} |\overline{v}(t) - \underline{v}(t)| + \underline{\sigma} \limsup_{t \to \infty} |\overline{v}(t) - \underline{v}(t)|
\end{align*}
\]
holds of all $x(0), \overline{v}(0), \underline{v}(0) \in \mathbb{R}^n$. Furthermore, if $\overline{w}_k, \underline{w}_k$ and $\overline{\xi}_k, \underline{\xi}_k$ are chosen to satisfy (6d) and $\overline{L}_k, \underline{L}_k$ are chosen for all $k = 1, 2, ..., m$, the implication
\[
\begin{align*}
\forall t \in [t_0, t] \quad \overline{v}(t) = \overline{v}(t) \Rightarrow \forall t \in [t_0, t] \quad \underline{v}(t) = \overline{v}(t)
\end{align*}
\]
holds true of all $x(0), \overline{v}(0), \underline{v}(0) \in \mathbb{R}^n$.

**Theorem 4.** Suppose that system (1a) is non-negative. If the matrices $A_k + 2L_k C$ and $A_k + 2L_k C$ are Metzler for all $k = 1, 2, ..., m$, then the system consisting of (1) and (5) equipped with (6a)-(6c) satisfies
\[
\begin{align*}
\forall t \in \mathbb{R}_+ \quad (\overline{v}(t) \leq x(t) \leq \underline{v}(t)) \\
\text{for all } x(0), \overline{v}(0) \in \mathbb{R}^n \text{ and } \underline{v}(0) \in \mathbb{R}^n \text{ satisfying} \\
(\overline{\sigma}) \leq (\underline{\sigma}) \leq 0.
\end{align*}
\]
\( \mu_k(x, u) = \tilde{\mu}_k(y, u), \quad \forall x \in \mathbb{R}^n, u \in \mathbb{R}^p \) \hspace{1cm} (17)

for all \( k = 1, 2, ..., m \), the use of \( \overline{\tau} \) and \( \underline{\tau} \) is not necessary. In fact, it is easy to see that replacing both \( \overline{\tau}_k \) and \( \underline{\tau}_k \) with \( \tilde{\mu}_k(y, u) \) in (5) yields
\[
\tilde{c} = \sum_{k=1}^{m} \tilde{\mu}_k(y, u)(A_k + T_k C)x + \tilde{d} - \tilde{d},
\]
\[
\tilde{\dot{x}} = \sum_{k=1}^{m} \tilde{\mu}_k(y, u)(A_k + L_k C)\tilde{x} + \tilde{d} - \tilde{d}.
\]

Therefore, all the statements of Theorems 3 and 4 hold true without the cause clause \( \lim_{t \to \infty} x(t) = 0 \) in (10).

4. RELAXATION OF ASSUMPTIONS

Theorem 4 requires the non-negativity of the plant (1a), and the Metzler property of \( A_k + T_k C \) and \( A_k + L_k C \). In order to relax these assumptions, one can confirm that the observer proposed in this paper allows one to use coordinate transformation which is popular in interval observer design (see, e.g., Raïssi et al. (2012); Efimov and Raïssi (2016)). In fact, applying \( z = Rx \) to the plant (1a) for a nonsingular \( R \in \mathbb{R}^{n \times n} \) yields
\[
\dot{\tilde{z}}(t) = \sum_{k=1}^{m} \tilde{\mu}_k(R^{-1}z(t), u(t))RA_kR^{-1}z(t) + R\beta(y(t), u(t)) + R\tilde{d}(t).
\]

The upper and lower bounds of disturbances are expressed in the new coordinate as
\[
R^+\tilde{\delta} - R^-\tilde{\delta}(t) \leq R\tilde{d}(t) \leq R^+\tilde{\delta} - R^-\tilde{\delta}.
\]

Then the non-negativity can be imposed on the system (18) with the transformed initial condition \( z(0) = Rx(0) \) and the transformed disturbance \( R\tilde{d}(t) \) instead of (1a). The matrix \( R \) provides a degree of freedom in securing the non-negativity for a given (1a). Using the disturbance bounds (19), the observer (5) can be built for \( z \) of (18) instead of \( x \). The properties in (6) for \( s = R^{-1}z \) can be defined by replacing \( \tau, \xi \) with \( \overline{\tau}, \underline{\tau} \), respectively. The initial conditions of the observer can be set to
\[
\overline{\tau}(0) = R^+\overline{\tau}(0) - R^-\overline{\tau}(0), \quad \overline{z}(0) = R^+\overline{z}(0) - R^-\overline{\tau}(0)
\]

since (14) implies \( \overline{z}(0) \leq z(0) \leq \overline{\tau}(0) \). Theorem 4 on the transformed coordinate requires that the transformed matrices \( R(A_k + T_k C)R^{-1} \) and \( R(A_k + L_k C)R^{-1} \) are Metzler, where \( R \) gives a degree of freedom. The interval estimation achieving (13) is obtained as
\[
\overline{\tau}(t) = S^+\overline{\tau}(t) - S^-\underline{\tau}(t), \quad \overline{z}(t) = S^+\overline{z}(t) - S^-\underline{\tau}(t),
\]

where \( S = R^{-1} \). Therefore, Theorem 4 holds true for \( z \) with the pair \( \overline{\tau}, \underline{\tau} \). It is important that (19) holds if and only if \( \overline{\tau} = \overline{\tau} \) (22) not only makes the interval estimation through \( \overline{\tau} \) and \( \underline{\tau} \) reasonable, but also allows (6d) to be effective with \( \overline{\tau} \) and \( \underline{\tau} \). Therefore, Theorem 3 is valid for the transformed variables \( z \) with \( \overline{\tau} \) and \( \underline{\tau} \) and it gives guaranteed properties on the original coordinate through (21). The transformed disturbance bounds (19) have a useful property that the two inequalities in (19) become equalities when \( \delta(t) = \delta(t) \).

5. AN EXAMPLE

Consider the plant (1) defined with
\[
A_1 = \begin{bmatrix} -2 & 2 \\ 5 & -6 \end{bmatrix}, \quad A_2 = 0, \quad A_3 = \begin{bmatrix} -2 & 0 \\ -4 & -2 \end{bmatrix} \quad (23a)
\]
\[
C = \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad \beta(y, u) = 0 \quad (23b)
\]
\[
\mu_1(x) = \frac{1}{2}, \quad \mu_2(x) = \frac{1}{2 + x_2}, \quad \mu_3(x) = \frac{x_2}{4 + 2x_2}, \quad (23c)
\]

which satisfy (2). Since (4) holds, this system (23) is non-negative. The plant (23) does satisfy the restrictive assumption made in the preliminary work of Ito and Dinh (2019) even if the disturbances are removed. By contrast, it can be verified that the plant (23) satisfies all the assumptions posed in Theorems 3 and 4 for
\[
L_1 = \begin{bmatrix} -2 \\ 0 \end{bmatrix}, \quad L_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad L_3 = \begin{bmatrix} -2 \\ 4 \end{bmatrix} \quad (24)
\]

with
\[
P = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}. \quad (25)
\]

All properties in (6) are achieved by
\[
\overline{w}_1 = w_1 = \frac{1}{2}, \quad \overline{w}_2 = \frac{1}{2 + x_2}, \quad \overline{w}_3 = \frac{1}{2 + \max(0, x_2)} \quad (26a)
\]
\[
\underline{w}_3 = \frac{\max(0, x_2)}{4 + 2x_2}, \quad w_3 = \frac{x_2}{4 + 2x_2}. \quad (26b)
\]

For the unmeasured state \( x_2 \), the interval computed by the proposed observer (5) is shown in Fig. 1 for
\[
x(0) = \begin{bmatrix} 3 \\ 0 \end{bmatrix}, \quad \tau(0) = \begin{bmatrix} 4 \\ 5 \end{bmatrix}, \quad \xi(0) = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \quad (27)
\]
\[
\delta(t) = \begin{bmatrix} \sin 4t + 1 \\ \frac{1 + t}{\cos 4t + 1} \end{bmatrix}, \quad \overline{\delta}(t) = \begin{bmatrix} \frac{1}{2} + t \\ \frac{1}{1 + t} \end{bmatrix}, \quad \delta(t) = 0. \quad (28)
\]

The interval estimation is also plotted for
\[
x(0) = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \quad \tau(0) = \xi(0) \quad (29)
\]
\[
\delta(t) = \begin{bmatrix} \sin 4t + 1 \\ \frac{1 + t}{\cos 4t + 1} \end{bmatrix}, \quad \overline{\delta}(t) = \delta(t) \quad (30)
\]

in Fig. 2. The simulation results are consistent with (10), (11), (12) and (13) established in Theorems 3 and 4. If one ignores (6d), all other conditions in (6) are met by
\[
\overline{w}_1 = \frac{1}{2}, \quad \overline{w}_1 = \frac{1}{2}, \quad \overline{w}_2 = \frac{1}{2}, \quad \overline{w}_3 = 0 \quad (31a)
\]
\[
\underline{w}_3 = \frac{1}{2}, \quad w_3 = 0. \quad (31b)
\]

This means that the observer (5) defined with (31) treats the nonlinearity of the unmeasured state \( x_2 \) as uncertainty, and the observer uses an upper bound and a lower bound of (23c). This treatment is basically the idea common in Chebotarev et al. (2015); Efimov and Raïssi (2016); Raïssi et al. (2012); Efimov et al. (2013). For the pair (27)-(28) and the pair (29)-(30), computed interval estimates are shown in Figs. 3 and 4, respectively. Comparing these plots with those in Figs. 1 and 2 confirms that the observer (5) with (26) achieving (6d) produces tighter interval estimates than the observer with (31) violating (6d).
Fig. 1. The interval $[x_2, \bar{x}_2]$ estimated by the observer (5) with (24) and (26) for plant (1) given by (23) with initial condition (27) and disturbance (28).

Fig. 2. The interval $[x_2, \bar{x}_2]$ estimated by the observer (5) with (24) and (26) for plant (1) given by (23) with initial condition (29) and disturbance (30); The three lines completely overlap.

Fig. 3. The interval $[x_2, \bar{x}_2]$ estimated by the observer (5) with (24) and (31) for plant (1) given by (23) with initial condition (27) and disturbance (28).

Fig. 4. The interval $[x_2, \bar{x}_2]$ estimated by the observer (5) with (24) and (31) for plant (1) given by (23) with initial condition (29) and disturbance (30).

6. CONCLUDING REMARKS

For interval observer design, this paper has proposed a mechanism of using estimated intervals to deal with nonlinearities of unmeasured variables. For polytopic models, it has been shown that the use of intervals can actually make the estimation better than treating the nonlinearities as uncertainty. The effectiveness has been demonstrated by a numerical example comparing two observers. In this paper, disturbances are addressed by both the observer structure and performance guarantees, which were not considered in the initial work presented in Ito and Dinh (2019) by the authors. Restrictive assumptions used in the initial work have also been removed, and such a point has also been illustrated by the numerical example.

Finally, it would be worth mentioning that it is possible to numerically reduce the $L_2$-gain from the disturbance to $\bar{x}$ as demonstrated in Efimov and Raïssi (2016). It is reasonable when $x$ is made small by control.

REFERENCES


Appendix A. APPENDIX

**Proof of Theorem 3 and Proposition 5**

Define

$$\overline{A}(\bar{x}, \bar{y}, u) = \sum_{k=1}^{m} \overline{p}_{k}(\bar{x}, \bar{y}, u) (A_k + L_k C)$$

$$\underline{A}(\underline{x}, \underline{y}, u) = \sum_{k=1}^{m} \underline{w}_{k}(\underline{x}, \underline{y}, u) (A_k + L_k C).$$

Then from (1a) and (5) we obtain

$$\dot{\bar{z}} = \underline{A}(\bar{x}, \bar{y}, u) \bar{z} + \overline{R}(x, \bar{x}, \bar{y}, u) x + \bar{z} - \delta$$

(A.1a)

$$\dot{\underline{z}} = \overline{A}(\underline{x}, \underline{y}, u) \underline{z} + \underline{R}(x, \underline{x}, \underline{y}, u) x + \underline{z} - \delta.$$  (A.1b)

The existence of $\bar{z} > 0$, $\underline{z} \geq 0$ and $\overline{P} > 0$ for $k = 1, 2, ..., m$ guarantees

$$\overline{P} \underline{A}(\bar{x}, \bar{y}, u) + \overline{A}(\bar{x}, \bar{y}, u) \overline{P} > 0$$

for all $\bar{x}, \bar{y} \in \mathbb{R}^n$ and $u \in \mathbb{R}^p$. Properties (7) and (9) yields

$$\overline{P} \underline{A}(\bar{x}, \bar{y}, u) + \underline{A}(\bar{x}, \bar{y}, u) \underline{P} + \overline{P} \overline{P} > 0.$$  (A.2)

The function $\nabla(\sigma) = \sigma^T \sigma$ satisfies

$$\nabla(\sigma) \leq -\frac{\sigma^T \sigma}{2} + \frac{2}{\sigma} \left( [R^T R] [R^T R] \right) \frac{x^T}{\sigma}$$

along (A.1a). From (15), for $\hat{z} = \sigma / 2$ we obtain

$$\nabla(\sigma) \leq -\frac{\sigma^T \sigma}{2} \min \left[ \begin{array}{c} 2 \left( x + 1 \right) \max \left[ \begin{array}{c} \frac{x^T}{\sigma} 
\end{array} \right] 
\end{array} \right]$$

where $\max$ (resp., $\min$) is the largest (resp., smallest) eigenvalue of $\sigma$. Hence, $\nabla(\sigma) = \sigma^T \sigma$ is an ISS Lyapunov function of system (A.1a) (see Sontag and Wang (1995)).

In the same way, $\nabla(\sigma) = \sigma^T \sigma$ is an ISS Lyapunov function of system (A.1b). Thereby, the implication (10) follows from ISS and (3). ISS of (A.1a) also gives

$$\limsup_{t \to \infty} |\sigma(t)| \leq \sqrt{\frac{2}{\sigma}} + \frac{\max}{\min} \limsup_{t \to \infty} \left[ \frac{x(t)}{d(t) - \delta(t)} \right]$$

Here, \(2 \frac{2}{\sigma} + \frac{\max}{\min} \limsup_{t \to \infty} \left[ \frac{x(t)}{d(t) - \delta(t)} \right]\) is the called the asymptotic gain. A similar asymptotic gain of (A.1b) is obtained between $\sigma^T \sigma \leq \delta^T \sigma$ and $\sigma(t)$. Since $\pi - x = \pi - \sigma$, and $||a, b||^2 \leq \pi^2 ||a, b||$ for $a, b \in R^n$, combining the two asymptotic gains yields (11) with

$\bar{g} = \frac{2}{\sigma} \sqrt{2} \min \left( \begin{array}{c} \frac{x^T}{\sigma} 
\end{array} \right), \quad \bar{g} = \frac{2}{\sigma} \sqrt{2} \min \left( \begin{array}{c} \frac{x^T}{\sigma} 
\end{array} \right)$

These arguments also give

$$\bar{g} = \frac{2}{\sigma} \sqrt{2} \min \left( \begin{array}{c} \frac{x^T}{\sigma} 
\end{array} \right), \quad \bar{g} = \frac{2}{\sigma} \sqrt{2} \min \left( \begin{array}{c} \frac{x^T}{\sigma} 
\end{array} \right)$$

if $\bar{g}(t) = \delta(t)$ holds for all $t \in R_+$. The definitions (A.2) and (A.3) prove the claims of Proposition 5. Finally, if (6d) holds, under the choice $L_k = L_k$ for all $k = 1, 2, ..., m$, summing up both sides of (A.1a) and (A.1b) yields, for all $x, \bar{x}, \underline{z} \in R^n$ and $u \in R^p$, $d(\bar{x} - \underline{z}) / dt \geq 0$ as long as $z = x$ and $\bar{z} \delta$. Hence, property (12) follows.

**Proof of Theorem 4**

The non-negativity assumption implies $x(t) \geq 0$ for all $t \in R_+$. By virtue of (6a) and (6b), the implications

$$\frac{x}{x} \leq x \leq \bar{x} \Rightarrow \bar{R}(x, \bar{x}, \bar{y}, u) \geq 0 \quad \frac{x}{x} \leq x \leq \bar{x} \Rightarrow \bar{R}(x, \bar{x}, \bar{y}, u) \geq 0 \quad \frac{x}{x} \geq 0 \text{ hold true for all } u \in R^p.$$  (A.4)

Due to property (6c), the matrix $\bar{R}(x, \bar{x}, \bar{y}, u)$ is Metzler for all $x, \bar{x}, \bar{y}$ and $u$ since $A_k + L_k C$ is Metzler for each $k = 1, 2, ..., m$. In the same way, $A_k + L_k C$ is Metzler. Suppose that there exists $i \in \{1, 2, ..., n\}$ such that $x(t_i) \geq L_i(t_i) = \bar{x}(t_i)$ and $x(t_i) \leq (t_i) = \pi(t_i)$ hold at some $t_i \in R_+$. Then (A.4) and $x(t_i) \geq 0$ imply $\bar{R}(x(t_i), \bar{x}(t_i), \pi(t_i), u(t_i)) x(t_i) \geq 0$. By virtue of (3), property (1a) yields $\bar{g}(t_i) \geq 0$. In the case where $x(t_i) = \pi(t_i)$ and $x(t_i) \leq \bar{x}(t_i)$ hold, we obtain $\bar{g}(t_i) \geq 0$ from (A.1b) and $x(t_i) \geq 0$ since (A.5) implies $\bar{R}(x(t_i), \bar{x}(t_i), \pi(t_i), u(t_i)) x(t_i) \geq 0$. Since the solutions $x, \bar{x}$ and $\bar{g}$ of (1a), (5a) and (5b) are continuously differentiable with respect to $t$, the restriction (14) of the initial condition results in (13).