Simultaneous state and unknown input set-valued observers for quadratically constrained nonlinear dynamical systems

Mohammad Khajenejad | Sze Zheng Yong

1Department of Mechanical and Aerospace Engineering, University of California, San Diego, California, USA
2School for Engineering of Matter, Transport and Energy, Arizona State University, Tempe, Arizona, USA

Correspondence
Sze Zheng Yong, School for Engineering of Matter, Transport and Energy, Arizona State University, Tempe, AZ 85287, USA. Email: szyong@asu.edu

Funding information
National Science Foundation grant CNS-1943545 and an Early Career Faculty grant 80NSSC21K0071 from NASA’s Space Technology Research Grants Program. This research was done while Mohammad Khajenejad was at Arizona State University.

Abstract
In this article, we propose fixed-order set-valued (in the form of $\ell_2$-norm hyperballs) observers for several classes of quadratically constrained nonlinear dynamical systems with unknown input signals that simultaneously/jointly find bounded hyperballs of states and unknown inputs that include the true states and inputs. Necessary and sufficient conditions in the form of linear matrix inequalities (LMIs) for the stability (in the sense of quadratic stability) of the proposed observers are derived for ($\mathcal{M}, \gamma$)-quadratically constrained ($\mathcal{M}, \gamma$)-QC systems, which includes several classes of nonlinear systems: (I) Lipschitz continuous, (II) ($\mathcal{A}, \gamma$)-QC* and (III) linear parameter-varying (LPV) systems. This new quadratic constraint property is at least as general as the incremental quadratic constraint property for nonlinear systems and is proven in the paper to embody a broad range of nonlinearities. In addition, we design the optimal $H_\infty$ observer among those that satisfy the quadratic stability conditions and show that the design results in uniformly bounded-input bounded-state (UBIBS) estimate radii/error dynamics and uniformly bounded sequences of the estimate radii. Furthermore, we provide closed-form upper bound sequences for the estimate radii and sufficient conditions for their convergence to steady state. Finally, the effectiveness of the proposed set-valued observers is demonstrated through illustrative examples, where we compare the performance of our observers with some existing observers.

KEYWORDS
nonlinear systems, resilient estimation, set-valued estimation, state and input estimation

1 | INTRODUCTION

1.1 | Motivation

Cyber-physical systems (CPS), for example, power grids, autonomous vehicles, medical devices, are systems in which computational and communication components are deeply intertwined and interacting with each other in several ways to control physical entities. Such safety-critical systems, if jeopardized or malfunctioning, can cause serious detriment to their operators, controlled physical components and the people utilizing them. A need for CPS security and for new designs of resilient estimation and control has been accentuated by recent incidents of attacks on CPS, for example, the Iranian nuclear plant, the Ukrainian power grid and the Maroochy water service. Specifically, false data injection attack...
is one of the most serious types of attacks on CPS, where malicious and/or strategic attackers inject counterfeit data signals into the sensor measurements and actuator signals to cause damage, steal energy and so forth. Given the strategic nature of these false data injection signals, they are not well-modeled by a zero-mean, Gaussian white noise nor by signals with known bounds. Nevertheless, reliable estimates of states and unknown inputs are indispensable and useful for the sake of attack identification, resilient control and so forth. Similar state and input estimation problems can be found across a wide range of disciplines, from input estimation in physiological systems, to fault detection and diagnosis, to the estimation of mean areal precipitation.

1.2 Literature review

Much of the research focus has been on simultaneous/joint input and state estimation for stochastic systems with unknown inputs, assuming that the noise signals are Gaussian and white, via minimum variance unbiased (MVU) estimation approaches (e.g., References 9-11), modified double-model adaptive estimation methods (e.g., Reference 12), or robust regularized least square approaches as in Reference 13. However, in order to address “set-membership” estimation problems in bounded-error settings, for example, in the presence of non-Gaussian or worst-case stochastic uncertainties with bounded support, as is considered in this article, where set-valued uncertainties are considered and sets of states and unknown inputs that are compatible with measurements are desired (cf. Reference 14 for a comprehensive discussion), the development of set-theoretic approaches are needed.

In the context of attack-resilient estimation against false data injection attacks, numerous approaches were proposed for deterministic systems (e.g., References 15-18), stochastic systems (e.g., References 19-21) and bounded-error systems (e.g., References 22-24). However, these approaches mainly yield point estimates, that is, the most likely or best single estimate, as opposed to set-valued estimates. On the other hand, the work in Reference 23 only computes error bounds for the initial state and Reference 22 assumes zero initial states and does not consider any optimality criteria.

In addition, unknown input observer designs for different classes of discrete-time nonlinear systems are relatively scarce. The method proposed in Reference 25 leverages discrete-time sliding mode observers for calculating state and unknown input point estimates, assuming that the unknown inputs have known bounds and evolve as known functions of states, which may not be directly applicable when considering adversaries in the system. The authors in Reference 26 proposed an LMI-based state estimation approach for globally Lipschitz nonlinear discrete-time systems, but did not consider unknown input reconstruction. An LMI-based approach was also used in Reference 27 for simultaneous/joint estimation of state and unknown input for a class of continuous-time dynamic systems with Lipschitz nonlinearities, but the authors did not address optimality nor stability properties for their observer, as well as only considered point estimates.

References 28 and 29 designed asymptotic observers to calculate point estimates for a class of continuous-time systems with bounded exogenous inputs, whose nonlinear terms satisfy an incremental quadratic inequality property, where the bounded unknown input assumption enables the authors to relax typically required observability/rank conditions. Moreover, in order to derive tractable LMIs, a specific subclass of incrementally quadratic constrained systems were considered. Similar work was done for the same class of discrete-time nonlinear systems in Reference 30, while the set-valued state estimation approach in Reference 31 uses mean value and first-order Taylor extensions to efficiently propagate constrained zonotopes through nonlinear mappings. However, none of them addressed unknown input estimation. Moreover, the restrictive assumption of bounded unknown inputs is needed in order to obtain convergent estimates.

Considering bounded unknown inputs, but with unknown bounds, the work in Reference 32 applied second-order series expansions to construct observer for state estimation in nonlinear discrete-time systems. The authors also provided sufficient conditions for stability and optimality of the designed estimator. However, their method does not compute unknown input estimates. On the other hand, in a recent and interesting work in Reference 33, the authors designed an observer for reconstruction of unknown exogenous inputs in nonlinear continuous-time systems with unknown and potentially unbounded inputs, providing sufficient LMI conditions for $\mathcal{L}_{\infty}$-stability of the observer. However, their observer does not simultaneously/jointly estimate the state, the unknown input estimates are point estimates and the optimality of their approach was not analyzed.

The author in Reference 14 and references therein discussed the advantages of set-valued observers (when compared to point estimators) in terms of providing hard accuracy bounds, which are important to guarantee
safety. In addition, the use of fixed-order set-valued methods can help decrease the complexity of optimal observers, which grows with time. Hence, a fixed-order set-valued observer for linear time-invariant discrete time systems with bounded errors, was presented in Reference 14, that simultaneously finds bounded hyperballs of compatible states and unknown inputs that are optimal in the minimum $H_\infty$-norm sense, that is, with minimum average power amplification. In our preliminary work, we extended the approach in Reference 14 to linear parameter-varying systems, while in Reference 37, we generalized the method to switched linear systems with unknown modes and sparse unknown inputs (attacks). In this work, we aim to further design novel hyperball-valued observers for a very broad class of quadratically constrained nonlinear systems, which can be considered as complementary to the work in References 38-46, where the authors considered state (and input) unknown modes and sparse unknown inputs (attacks). In this work, we aim to further design novel hyperball-valued observers of compatible states and unknown inputs that are optimal in the minimum $H_\infty$-norm sense, that is, with minimum average power amplification.

First, we introduce a novel class of time-varying nonlinear vector fields that we call $(\mathcal{M}, \gamma)$-Quadratically Constrained ($(\mathcal{M}, \gamma)$-QC) functions and show that they include a broad range of nonlinearities. We also derive some results on the relationship between $(\mathcal{M}, \gamma)$-QC functions with other classes of nonlinearities, such as the incrementally quadratically constrained, Lipschitz continuous and linear parameter-varying (LPV) functions.

Then, we present our three-step recursive set-valued observer for nonlinear discrete-time systems. In particular, in the absence of noise, we derive sufficient and necessary conditions in the form of LMIs for the existence and stability of the observer in the sense of quadratic stability for this class of $(\mathcal{M}, \gamma)$-QC functions, as well as three classes of nonlinearities: (I) Lipschitz continuous, (II) $(A, \gamma)$-QC*, and (III) LPV systems.

Furthermore, we design $H_\infty$ observers among those that satisfy the quadratic stability conditions using semi-definite programs with additional LMIs constraints for each of the aforementioned system classes. Then, we show that our $H_\infty$ observer design leads to estimate radii dynamics that are uniformly bounded-input bounded-state (UBIBS), which equivalently results in uniformly bounded sequences of the estimate radii in the presence of noise. Moreover, we derive closed-form expressions for upper bound sequences for the estimate radii as well as sufficient conditions for the convergence of these radii upper bound sequences to steady state.

Note that we consider completely unknown inputs (different from noise signals) without imposing any assumptions on them (such as being norm bounded, with limited energy or being included in some known set). Considering resilient estimation in cyber-physical systems, our set-valued observers are applicable for achieving attack-resiliency against false data injection attacks on both actuator and sensor signals. It is worth mentioning that in our preliminary work, we designed hyperball-valued $H_\infty$ observers for the special case of LPV systems.

### 1.3 Contribution

The goal of this article is to bridge the gap between set-valued state estimation without unknown inputs and point-valued state and unknown input estimation for a broad range of time-varying nonlinear dynamical systems with nonlinear observation functions. In particular, we propose fixed-order set-valued (in the form of $\ell_2$-norm hyperballs) observers for nonlinear discrete-time bounded-error systems that simultaneously/jointly find uniformly bounded sets of states and unknown inputs that contain the true state and unknown input, are compatible/consistent with measurement outputs as well as the nonlinear observation functions, and are optimal in the minimum $H_\infty$-norm sense, that is, with minimum average power amplification.

### 1.4 Notation

$\mathbb{R}^n$ denotes the $n$-dimensional Euclidean space, $\mathbb{C}$ is the set of complex numbers, $\mathbb{Z}$ nonnegative integers and $\mathbb{N}$ positive integers, while $\mathbb{R}$, $\mathbb{R}_+$ and $\mathbb{R}_{++}$ denote the sets of real, non-negative and positive real numbers, respectively, and $0_{n \times m}$ denotes a zero matrix in $\mathbb{R}^{n \times m}$. For a vector $v \triangleq [v_1, \ldots, v_n]^\top \in \mathbb{R}^n$ and a matrix $M \in \mathbb{R}^{p \times q}$, $\|v\| \triangleq \sqrt{v^\top v}$ and $\|M\|$ denote their (induced) 2-norm, while $\|v\|_\infty \triangleq \max_{1 \leq e \leq \ldots, n} |v_e|$ denotes the $\infty$-norm of $v$. Moreover, the transpose, inverse, Moore–Penrose pseudoinverse and rank of $M$ are given by $M^\top$, $M^{-1}$, $M^\dagger$, and $\text{rk}(M)$. For symmetric matrices $S$ and $S'$, $S > 0$, $S \geq 0$, $S < 0$, and $S \leq 0$ mean that $S$ is positive semidefinite, positive definite, positive seminegative and positive negative, respectively. Moreover, $S \geq S'$ and $S \leq S'$ mean $S - S'$ is positive semidefinite and negative semidefinite, respectively.
1.5 Preliminary material

In this subsection, we state a prior result on affine abstraction of nonlinear systems, which will be used in the next section.

**Proposition 1** (Affine abstraction\textsuperscript{47}). Consider the vector field \( q(\cdot) : B \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m \), where \( B \triangleq [\underline{z}, \overline{z}] \) is an \( n_q \)-dimensional hyper-interval with \( \underline{z}, \overline{z} \) being its maximal and minimal values, respectively. Suppose \( \overline{A}_n, A_n, \overline{e}_n, \underline{e}_n, \theta_n \) is a solution of the following linear program (LP):

\[
\min_{\theta} \theta \\
\text{s.t. } \left( \overline{A} - A \right) z + \overline{e} - \theta \leq f_k - \underline{e}, \quad \forall z \in V_B,
\]

where \( V_B \) is the vertex set of \( B \), \( 1_m \in \mathbb{R}^m \) is a vector of ones and \( \sigma \) can be computed via Reference \textsuperscript{47} (proposition 1) for different function classes. Then, \( \overline{A}_n z + \overline{e} \leq q(z) \leq \overline{A}_n z + \overline{e}, \forall z \in B \). We call \( \overline{A}_n \) upper and lower affine abstraction gradients of function \( q(\cdot) \) on \( B \).

**Remark 1.** The main idea in (1) is to find the two (tightest) upper and lower affine hyperplanes that bound/frame/abstract the nonlinear function/vector field \( q(\cdot) \) for all the evaluations of \( q(\cdot) \) in the hyper-interval \( B \). It is worth noting that all the inequalities in (1) are component-wise. Moreover, it is worth emphasizing that \( q(z) \) is the evaluation of the nonlinear function \( q(\cdot) \) at the known vertex point \( z \in V_B \), where \( V_B \) is the vertex set of \( B \), and thus, is a constant. Consequently, the problem in (1), is always linear with respect to all its variables \( \theta, \overline{A}_n, A_n, \overline{e}_n, \underline{e}_n \), that is, is an LP regardless of \( q(\cdot) \), and hence, can be easily solved using off-the-shelf solvers, for example, Gurobi\textsuperscript{48}. Interested readers are referred to Reference \textsuperscript{47} for more details.

2 PROBLEM STATEMENT

2.1 Problem formulation

In this section, we describe the system, vector field and unknown input signal assumptions as well as formally state the observer design problem.

**System assumptions.** Consider the following nonlinear time-varying discrete-time bounded-error system

\[
x_{k+1} = f_k(x_k) + B_k u_k + G d_k + W w_k, \\
y_k = C x_k + D_k u_k + H d_k + v_k,
\]

where \( x_k \in \mathcal{X} \subseteq \mathbb{R}^n \) is the state vector at time \( k \in \mathbb{N} \), \( u_k \in \mathcal{U} \subseteq \mathbb{R}^m \) is a known input vector and \( y_k \in \mathbb{R}^l \) is the measurement vector. The process noise \( w_k \in \mathbb{R}^n \) and the measurement noise \( v_k \in \mathbb{R}^l \) are assumed to be bounded (and non-Gaussian), with \( \|w_k\| \leq \eta_w \) and \( \|v_k\| \leq \eta_v \) (thus, they are \( \ell_\infty \) sequences) and \( W \) is known and of appropriate dimension. Note that considering matrix \( W \) allows us to only impose noise on a subset of the states, which is very common for dynamical systems with inertia, such as a car/vehicle dynamics. We also assume an estimate \( \hat{x}_0 \) of the initial state \( x_0 \) is available, where \( \|\hat{x}_0 - x_0\| \leq \delta_0^x \), with known \( \delta_0^x \in \mathbb{R}_+ \).

The mapping \( f_k(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^n \) is a known time-varying nonlinear function, while \( d_k \in \mathbb{R}^p \) is an arbitrary unknown input signal that affect the state and observation equations through the known matrices \( G \in \mathbb{R}^{m \times p} \) and \( H \in \mathbb{R}^{l \times p} \). Furthermore, \( B_k \in \mathbb{R}^{m \times n} \) and \( D_k \in \mathbb{R}^{l \times m} \) are known time-varying matrices at each time step \( k \). Note that without loss of generality, we assume that \( \text{rk}[G^T H^T] = p, n \geq l \geq 1, l \geq p \geq 0, \) and \( m \geq 0 \).

**Unknown input (or attack) signal assumptions.** The unknown input signal \( d_k \) is not constrained to be a signal of any type (random or strategic) nor to follow any model, thus no prior “useful” knowledge or information of the dynamics or bounds of \( d_k \) is available (independent of \( \{d_k\} \forall k \neq \ell \), \( \{v_k\} \) and \( \{v_k\} \forall \ell \)). Consequently, the bounds and the (possibly non-Gaussian) distribution of \( d_k \) are unknown, and thus, \( d_k \) is suitable for representing adversarial attack signals.

**Vector field assumptions.** Here, we formally state the classes of nonlinear systems, related to the assumptions about the nonlinear, time-varying vector field \( f_k(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^n = \begin{bmatrix} f_{k,1}(\cdot) & \cdots & f_{k,j}(\cdot) & \cdots & f_{k,n}(\cdot) \end{bmatrix}^T, \forall j \in \{1, \ldots, n\}, \forall k \in \mathbb{Z} \), that
we consider in this article. In particular, we consider the following classes of nonlinear systems, which will be formally defined in Section 3:

**Class 0.** $(\mathcal{M}, \gamma)$-QC systems, with some known $\mathcal{M} \in \mathbb{R}^{2n \times 2n}$ and $\gamma \in \mathbb{R}_+$.  

**Class I.** Globally $L_f$-Lipschitz continuous systems.  

**Class II.** $(\mathcal{A}, \gamma)$-QC* systems, with some known $\mathcal{A} \in \mathbb{R}^{n \times n}$ and $\gamma \in \mathbb{R}_+$.  

**Class III.** LPV systems with constituent matrices $A^i \in \mathbb{R}^{n \times n}$, $i \in \{1, \ldots, N\}$.  

For Class III of systems, the system dynamics is governed by an LPV system with known parameters at run-time. We call each tuple $(A^i, C, G, H), i \in \{1 \ldots N\}$, an LTI constituent of system (2).

The simultaneous/joint input and state set-valued observer design problem is twofold and can be stated as follows:

**Problem 1.** Given the nonlinear discrete-time bounded-error system with unknown inputs (2) (cf. Remark 2),

1. Design stable observers that simultaneously/jointly find bounded sets of compatible states and unknown inputs for the four classes of nonlinear systems.
2. Among the observers that satisfy 1, find the optimal observer in the minimum $H_{\infty}$-norm sense, that is, with minimum average power amplification.

## 2.2 Extension to more general cases

It is noteworthy that System (2) can be easily extended in several ways to cover much more general classes of nonlinear dynamics, for example, to include the case where different attack signals are injected onto the sensors and actuators as well as the case where the attack signals compromise the system in a nonlinear manner. To illustrate this, consider the following dynamical system:

$$
\begin{align*}
x_{k+1} &= f_k(x_k) + B_k \hat{u}_k + \tilde{G} \tilde{h}_k(x_k, \hat{u}_k, \delta_k) + W_k, \\
y_k &= \mu(x_k, \hat{u}_k) + \tilde{D} \tilde{h}_k(x_k, \hat{u}_k, \delta_k) + \tilde{v}_k,
\end{align*}
$$

which is an extension of System (2), where $\hat{u}_k \in \mathbb{R}^m$ is the known input, $\delta_k \in \mathbb{R}^p$ and $\delta_k \in \mathbb{R}^q$ can be interpreted as arbitrary (and different) unknown inputs that affect the state and observation equations through the known time-varying nonlinear vector fields $\tilde{G} \tilde{h}_k(\ldots) : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{p_1} \to \mathbb{R}^{n_1}$ and $\tilde{h}_k(\ldots) : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{p_2} \to \mathbb{R}^{n_2}$, respectively. Moreover, $\tilde{G} \in \mathbb{R}^{n \times n}$ and $\tilde{h} \in \mathbb{R}^{l \times m}$ are known time-invariant matrices, whereas $B_k \in \mathbb{R}^{n \times m}$ and $D_k \in \mathbb{R}^{l \times m}$ are known time-varying matrices at each time step $k$.

On the other hand, $\mu(\cdot, \cdot) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^l$ is a known observation mapping for which we consider two cases:

**Case 1.** $\mu(x_k, \hat{u}_k) = C x_k + D \hat{u}_k$, that is, $\mu(\cdot, \cdot)$ is linear in $x_k$ and $\hat{u}_k$.

**Case 2.** $\mu(\cdot, \cdot)$ is nonlinear with bounded interval domains, that is, there exist known intervals $\mathcal{X}$ and $\mathcal{U}$ such that $\mathcal{X} \subseteq \mathcal{X} \subseteq \mathbb{R}^n$ and $\mathcal{U} \subseteq \mathcal{U} \subseteq \mathbb{R}^m$.

In the second case, we can apply our previously developed **affine abstraction** tools in Reference 47 (cf. Proposition 1) to derive affine upper and lower abstractions for $\mu(\cdot, \cdot)$ using Proposition 1 and the linear program therein to obtain $\overline{C}, \underline{C}, \overline{D}, \underline{D}, \overline{\epsilon}$ and $\underline{\epsilon}$ with appropriate dimensions, such that for all $x_k \in \mathcal{X}$ and $\hat{u}_k \in \mathcal{U}$:

$$
\begin{align*}
\overline{C} x_k + \overline{D} \hat{u}_k + \overline{\epsilon} &\leq \mu(x_k, \hat{u}_k) \\
\underline{C} x_k + \underline{D} \hat{u}_k + \underline{\epsilon} &\leq \overline{C} x_k + \overline{D} \hat{u}_k + \overline{\epsilon},
\end{align*}
$$

where “$\leq$” holds in a component-wise manner. Next, by taking the average of the upper and lower affine abstractions in (4) and adding an additional bounded disturbance/perturbation term $v_k$ (with its $\infty$-norm being less than half of the maximum distance), it is straightforward to reformulate the inequalities in (4) as the following equality:
with $C \triangleq \frac{1}{2}(\overline{C} + C)$, $D \triangleq \frac{1}{2}(\overline{D} + D)$, $e \triangleq \frac{1}{2}(\overline{e} + e)$, $v_k^e \sqsubseteq \eta^e \triangleq \frac{1}{2}\theta^e$, where $\theta^e$ is the solution to the LP in (1). In other words, the equality in (5) is a “redefinition” of the inequalities in (4), which is obtained by adding the uncertain noise $v_k^e$ to the midpoint (center) of the interval in (4) (i.e., $C_{x_k} + \overline{D}u_k + e = \frac{1}{2}(C_{x_k} + \bar{D}u_k + e + C_{x_k} + \overline{D}u_k + \bar{e})$, to recover all possible $\mu(x_k, \hat{u}_k)$ in the interval given by (5). In a nutshell, the above procedure “approximates” $\mu(x_k, \hat{u}_k)$ with an appropriate linear term and accounts for the “approximation error” using an additional disturbance/noise term.

Then, using (5), the system in (3) can be rewritten as:

$$
\begin{align*}
  x_{k+1} &= f_k(x_k) + B_k u_k + \hat{G}g_k(x_k, u_k, d_k^o) + Ww_k, \\
  y_k &= Cx_k + D_k u_k + \hat{H}h_k(x_k, u_k, d_k^o) + v_k,
\end{align*}
$$

where $B_k \triangleq \hat{B}_k 0]$, $u_k \triangleq [\hat{u}_k^T \ e^T]^T$, $g_k(x_k, u_k, d_k^o) \triangleq \hat{g}_k(x_k, \hat{u}_k, d_k^o)$, $D_k \triangleq [\overline{D}_k \ D_k I]$, $h_k(x_k, u_k, d_k^o) \triangleq \hat{h}_k(x_k, \hat{u}_k, d_k^o)$, and $v_k = \hat{v}_k + v_k^e$ with $\|v_k\| \leq \eta_0 \triangleq \eta_0 + \eta_0^e$. Similarly, it is straightforward to notice that Case 1 can also be represented by (6) with $B_k \triangleq \hat{B}_k$, $u_k \triangleq \hat{u}_k$, $D_k \triangleq \overline{D}_k + D$ and $v_k \triangleq \hat{v}_k$ with $\|v_k\| \leq \eta_0 \triangleq \eta_0^e$.

Now, courtesy of the fact that the unknown input signals $d_k^o$ and $d_k^o$ in (6) can be completely arbitrary, by lumping the nonlinear functions with the unknown inputs in (6) into a newly defined unknown input signal $d_k \triangleq [g_k(x_k, u_k, d_k^o) \ h_k(x_k, u_k, d_k^o)] \in \mathbb{R}^{p}$, as well as defining $G \triangleq [\hat{G} \ 0_{nxn_\eta}], H \triangleq [\hat{0}_{nxn_\eta} \ \hat{H}]$, we can equivalently transform system (6) to a new representation, precisely in the form of (2).

Remark 2. From the discussion above, we can conclude that set-valued state and input observer designs for system (2) are also applicable to system (3), with the slight difference in input estimates that the latter returns set-valued estimates for $d_k \triangleq \left[ \begin{array}{c} g_k(x_k, u_k, d_k^o) \\ h_k(x_k, u_k, d_k^o) \end{array} \right]$, where we can apply any preimage set computation techniques in the literature such as References 49-51 to find set estimates for $d_k^o$ and $d_k^o$ using the set-valued estimate for $x_k$ and the known $u_k$. Given this, throughout the rest of the article, we will consider the design of set-valued state and unknown input observers for system (2) (with sets in the form of hyperballs).

Remark 3. Note that the case where the feedthrough matrix in System (2) is zero, that is, $H = 0$, as well as the case where the process and sensors in (2) are degraded by different attack (unknown input) signals, are both special cases of the system (3), where $\hat{H} = 0$, $\hat{g}$, $\hat{h}$ are affine functions, respectively; thus, these cases can also be considered with our proposed framework.

### 3 STRUCTURAL PROPERTIES

Here, we briefly introduce the structural properties that we will consider for the aforementioned system classes 0, I, II–III, so that we will be able to refer to them later when needed.

**Definition 1 (Strong detectability):** The following bounded-error linear time invariant (LTI) system:

$$
\begin{align*}
  x_{k+1} &= Ax_k + Bu_k + Gd_k + w_k, \\
  y_k &= Cx_k + Du_k + Hd_k + v_k,
\end{align*}
$$

that is, the quadruple $(A, G, C, H)$, is strongly detectable if $y_k = 0 \ \forall \ k \geq 0$ implies $x_k \to 0$ as $k \to \infty$, for all initial states and input sequences $\{d_i\}_{i \in \mathbb{N}}$, where $A, B, G, C, D, H$ are known constant matrices with appropriate dimensions, and $x_k$, $u_k$, $y_k$, $d_k$, $w_k$, and $v_k$ are system state, known input, output, unknown input, bounded norm process noise, and measurement noise signals, respectively.

**Remark 4.** Several necessary and sufficient rank conditions are provided in Reference 14 (theorem 1) to check the strong detectability of system (7), that is, $(A, G, C, H)$, including rk $R_z(z) \triangleq \text{rk} \begin{bmatrix} zI - A & -G \\ C & H \end{bmatrix} = n + p$ for all $z \in \mathbb{C}, |z| \geq 1$. It is worth mentioning that all the aforementioned conditions are equivalent to the system being minimum-phase (i.e., the
invariant zeros of $R_s(z)$ are stable). Moreover, strong detectability implies that the pair $(A, C)$ is detectable, and if $l = p$, then strong detectability implies that the pair $(A, G)$ is stabilizable (cf. Reference 14 (theorem 1) for more details).

**Definition 5** (Uniform global Lipschitzness). A time-varying vector field $f_k(\cdot) : D_k \to \mathbb{R}^m$ is uniformly globally $L_f$-Lipschitz continuous on $D_k \subseteq \mathbb{R}^n$, if there exists a time independent $L_f \in \mathbb{R}_{++}$, such that for any time step $k \in \mathbb{Z}$, $\|f_k(x_k^1) - f_k(x_k^2)\| \leq L_f \|x_k^1 - x_k^2\|$, $\forall x_k^1, x_k^2 \in D_k$. In other words, $f_k$ is globally $L_f$-Lipschitz, uniformly in time.

**Definition 3** (Time-varying LPV functions). A time-varying vector field $f_k(\cdot) : \mathbb{R}^p \to \mathbb{R}^q$ is linear parameter-varying (LPV), if at each time step $k \in \mathbb{Z}$, $f_k(x_k)$ can be decomposed into a convex combination of linear functions with known coefficients, that is, $\forall k \geq 0$, $\exists N \in \mathbb{N}$ such that $\forall i \in \{1, \ldots, N\}$, there exist known, time-varying $\lambda_{i,k} \in [0, 1]$ and $A_i \in \mathbb{R}^{pq}$ such that $\sum_{i=1}^N \lambda_{i,k} = 1$ and $f_k(x_k) = \sum_{i=1}^N \lambda_{i,k} A_i x_k$. Each linear function $A_i x_k$ is called a constituent function of the original nonlinear time-varying vector field $f_k(x_k)$ at $x_k \in D_k$.

Next, through the following definition, we slightly generalize the class of $\delta$-QC functions introduced in Reference 28 to time-varying $\delta$-QC vector fields.

**Definition 4** (Time-varying $\delta$-QC mappings). A symmetric matrix $M \in \mathbb{R}^{(n_q+n_l)\times(n_q+n_l)}$ is an incremental multiplier matrix ($\delta$MM) for time-varying vector field $f_k(\cdot) \in D_k \subseteq \mathbb{R}^n$, if the following incremental quadratic constraint ($\delta$QC) is satisfied for all $q_1, q_2 \in D_k$ and for all time steps $k \in \mathbb{Z}$: $[\Delta f_k^T \ Delta q] M [\Delta f_k^T \ Delta q]^T \geq 0$, where $\Delta q \triangleq q_2 - q_1$ and $\Delta f_k \triangleq f_k(q_2) - f_k(q_1)$.

Now, we introduce a new class of systems we call $(M, \gamma)$- quadratically constrained $(\mathcal{M}, \gamma)$-QC systems that is at least as general as $\delta$-QC systems and includes a broad range of nonlinearities.

**Definition 5** (Time-varying $(M, \gamma)$-QC functions). A time-varying vector field $f_k(\cdot) : D_k \subseteq \mathbb{R}^p \to \mathbb{R}^q$ is $(M, \gamma)$- Quadratically Constrained, that is, $(M, \gamma)$-QC, if there exist symmetric matrix $\mathcal{M} \in \mathbb{R}^{(p+q)\times(p+q)}$ and $\gamma \in \mathbb{R}$ such that

$$[\Delta f_k^T \ Delta x] \mathcal{M} [\Delta f_k^T \ Delta x]^T \geq \gamma,$$

for all $x_1, x_2 \in D_k$ and for all time steps $k \in \mathbb{Z}$, where $\Delta x \triangleq x_2 - x_1$ and $\Delta f_k \triangleq f_k(x_2) - f_k(x_1)$. We call $\mathcal{M}$ a multiplier matrix for function $f_k(\cdot)$.

First of all, we show that a vector field may satisfy $(M, \gamma)$-QC property with different pairs of $(M, \gamma)$'s. For clarity, all proofs are provided in the Appendix.

**Proposition 2.** Suppose $f_k(\cdot)$ is $(M, \gamma)$-QC. Then it is also $(\kappa M, \kappa \gamma)$-QC, $(\nu M, \nu \gamma)$-QC, $(M, \rho)$-QC and $(M', \gamma)$-QC for every $\kappa \geq 0, \nu \geq 1, \rho \leq \gamma$ and $M' \supseteq M$.

It is worth mentioning that $(M, \gamma)$-QC is a very broad class of systems. In particular, we next show in Propositions 3 and 4 that the $(M, \gamma)$-QC property includes monotonicity and Lipschitz continuity, respectively, which themselves are broad classes of systems. Moreover, as will be shown in Proposition 5, $(M, \gamma)$-QC is at least as general as the incremental quadratic constraint ($\delta$QC) property (cf. Definition 4), which recently has received considerable attention in nonlinear system state and input estimation, and as shown in References 28,30,33, is a superset of several classes of nonlinearities, for example, Lipschitz continuous, matrix parametrized and incrementally sector bounded systems (cf. Figure 1). Consequently, the class of $(M, \gamma)$-QC functions is a generalization of several types of nonlinearities (cf. Corollary 1 and Figure 1) and hence, covers a very wide range of nonlinear systems.

**Proposition 3.** Suppose the vector field $f_k : D_k \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is monotone nondecreasing, that is, $\Delta f_k^T \Delta x_k \geq 0$, for all $x_1, x_2 \in D_k$, or is monotonically nondecreasing, that is, $\Delta f_k^T \Delta x_k \leq 0$, for all $x_1, x_2 \in D_k$, in its domain $D_k$, where $\Delta f_k$ and $\Delta x_k$ are defined in Definition 5. Then, $f_k$ is $\delta$-QC and $(M, \gamma)$-QC with $M = M = \begin{bmatrix} 0_{n \times n} & I_{n \times n} \\ I_{n \times n} & 0_{n \times n} \end{bmatrix}$ and any $\gamma \geq 0$, or $\delta$-QC and $(-M, \gamma)$-QC with $M = M = \begin{bmatrix} 0_{n \times n} & I_{n \times n} \\ I_{n \times n} & 0_{n \times n} \end{bmatrix}$ and any $\gamma \leq 0$, respectively.

**Proposition 4.** Every globally $L_f$-Lipschitz continuous function is $\delta$-QC with multiplier matrix $M = \begin{bmatrix} -I & 0 \\ 0 & L_f^2 \end{bmatrix}$, but the converse is not true, that is, there exist $\delta$-QC functions that are not globally Lipschitz continuous. In other words, the class of globally Lipschitz continuous functions is a strict subset of the class of $\delta$-QC functions.

**Proposition 5.** Every nonlinearity that is $\delta$-QC with multiplier matrix $M$ is $(M, \gamma)$-QC for any $\gamma \leq 0$. 
Proposition 7. Suppose $f_k(\cdot)$ is globally $L_f$-Lipschitz continuous and the state space, $\mathcal{X}$, is bounded, that is, there exists $r \in \mathbb{R}_+$ such that for all $x \in \mathcal{X}$, $\|x\| \leq r$. Then, $f_k(\cdot)$ is a $(A, \gamma)$-QC* function with $A = 0_{n \times n}$ and $\gamma = -4r^2 L_f^2$.

Proposition 7. Suppose $f_k(\cdot)$ is $(A, \gamma)$-QC* with some $\gamma \geq 0$ and $A \neq 0_{n \times n}$. Then, $f_k(\cdot)$ is globally $L_f$-Lipschitz continuous with $L_f = \sqrt{\lambda_{\max}(A^T A)}$.

Lemma 1. Suppose vector field $f_k(\cdot)$ can be decomposed as the sum of an affine and a bounded nonlinear function $g_k(\cdot)$ at each time step $k \in \mathbb{Z}$, that is, $f_k(x) = Ax + h + g_k(x)$, $\forall k \in \mathbb{Z}$, where $A \in \mathbb{R}^{n \times n}$, $h \in \mathbb{R}^n$ and $\|g_k(x)\| \leq r \in \mathbb{R}_+$ for all $x \in D_{\mathcal{F}}$ and all $k \in \mathbb{Z}$. Then, $f_k(\cdot)$ is an $(A, \gamma)$-QC* function with $A = A$ and any $\gamma \leq -(2r)^2$. 

Corollary 1. Incrementally sector bounded nonlinearities and nonlinearities with matrix parameterizations and so forth, which are $\delta$-QC (cf. Figure 1 and Reference 28 (sections 5.1–5.2)), are also $(\mathcal{M}, \gamma)$-QC (the reader is referred to References 28, 30, 33 for definitions, demonstrations and more detailed examples). Moreover, the set of globally Lipschitz nonlinear functions is a strict subset of $(\mathcal{M}, \gamma)$-QC functions.

Next, we provide some instances of nonlinear $(\mathcal{M}, \gamma)$-QC vector fields, where we show that the corresponding $\mathcal{M}$ and $\gamma$ can be easily found/computed, whereas in general, it is difficult/unclear how to find the multiplier $M$ matrix, even if these were $\delta$-QC vector fields.

Example 1. Consider any monotone (nondecreasing or nonincreasing) vector filed such as $g(x) = x^5$ with $D_g = \mathbb{R}$ or $h(x) = \tan(x)$ with $D_h = (-\frac{\pi}{2}, \frac{\pi}{2})$. By Proposition 3, $g$ and $h$ are $(\mathcal{M}, \gamma)$-QC with $\mathcal{M} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and any $\gamma \geq 0$.

Example 2. Now, consider $f(x) = x^2$ with $D_f = [-\bar{x}, \bar{x}] \in \mathbb{R}$, $\bar{x} \geq 0.5$, which is not a monotone function. Let $\mathcal{M}_0 = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$. It can be verified that $\begin{bmatrix} (\Delta f)^T \\ (\Delta x)^T \end{bmatrix} \mathcal{M}_0 \begin{bmatrix} (\Delta f)^T \\ (\Delta x)^T \end{bmatrix}^T = \|\Delta f - \Delta x\|^2 = \|(\Delta x)^2 - \Delta x\|^2 \geq -[2\bar{x}(2\bar{x} + 1)]^2 = -9$, for $x_1, x_2 \in D_f$. Hence, $f(x) = x^2$ for all $x \in [-\bar{x}, \bar{x}] \in \mathbb{R}$ with $\bar{x} \geq 0.5$ is $(\mathcal{M}_0, -9)$-QC.

Furthermore, considering a specific structure for the multiplier matrix $\mathcal{M}$, we introduce a new class of functions that is a subset of the $(\mathcal{M}, \gamma)$-QC class.

Definition 6 (Time-varying $(A, \gamma)$-QC* functions). A time-varying vector field $f_k(\cdot)$ is an $(A, \gamma)$-QC* function, if it is $(\mathcal{M}, \gamma)$-QC for some $\gamma \in \mathbb{R}$ and there exists a known $A \in \mathbb{R}^{n \times n}$, such that $\mathcal{M} = \begin{bmatrix} -I_{n \times n} & A \\ A^T & -A^T A \end{bmatrix}$. We call $A$ an auxiliary multiplier matrix for function $f_k(\cdot)$.

Now we present some results that establish the relationships between the aforementioned classes of nonlinearities.
Note that some \((\mathcal{M}, \gamma)\)-QC systems are also \((A, \gamma)\)-QC\(^*\). The following Proposition 8 helps with finding such an \(A\) for some specific structures of \(\mathcal{M}\).

**Proposition 8.** Suppose \(f_k(\cdot) : \mathbb{R}^{2n} \to \mathbb{R}^{2n}\) is an \((\mathcal{M}, \gamma)\)-QC vector field, with \(\mathcal{M} = \begin{bmatrix} M_{11} & M_{12} \\ M_{12}^T & M_{22} \end{bmatrix}\), where \(M_{11}, M_{12}, M_{22} \in \mathbb{R}^{p \times n}\), \(M_{11} + I_{n \times n} \leq 0\) and \(M_{22} + M_{12}^T M_{12} \leq 0\). Then, \(f_k(\cdot)\) is an \((A, \gamma)\)-QC\(^*\) function with \(A = M_{12}\).

The reader can verify that such sufficient conditions in Proposition 8 hold for the function in Example 2.

**Proposition 9.** Every LPV function \(f_k(\cdot)\) with constituent matrices \(A_i, \forall i \in 1 \ldots N\), is \(\|A^n\|\)-globally Lipschitz continuous, where \(\|A^n\| = \max_{i \in 1 \ldots N} \|A_i\|\).

**Corollary 2.** As a direct corollary of Propositions 6 and 9, any bounded domain LPV function is an \((A, \gamma)\)-QC\(^*\) function.

Figure 1 summarizes all the above results on the relationships between several classes of nonlinearities.

### 4 | Fixed-Order Simultaneous Input and State Set-Valued Observer Framework

In this section we introduce our observer structure for simultaneous/joint input and state estimation and its properties. Specifically, we propose recursive set-valued observers that consist of three steps: (i) an unknown input estimation step that returns the set of compatible unknown inputs using the current measurement and the set of compatible states, (ii) a time update step in which the compatible set of states is propagated based on the system dynamics, and (iii) a measurement update step where the set of compatible states is updated according to the current measurement. Since the complexity of optimal observers increases with time, we will only focus on fixed-order recursive filters, similar to References 14, 34, 52, and in particular, we consider set-valued estimates in the form of hyperballs. Before mathematically introducing the proposed observer, we first provide a formal definition of a hyperball in the Euclidean space.

**Definition 7** (Hyperball). A hyperball in \(\mathbb{R}^n\) with radius \(r \in \mathbb{R}_+\) and center/centroid \(c \in \mathbb{R}^n\), \(B_n(r, c)\), is the set of all points with 2-norm distances that are not greater than \(r\) from \(c\), that is,

\[
B_n(r, c) \equiv \{ x \in \mathbb{R}^n \ | \ |x - c| \leq r \}.
\]

Given Definition 7, in this work, our set-valued estimates are characterized as follows:

\[
\hat{d}_{k-1} = \{ d \in \mathbb{R}^p : |d_{k-1} - \hat{d}_{k-1}| \leq \delta_{d_{k-1}} \},
\]

\[
\hat{\delta}_{\hat{x}_{k-1}} = \{ x \in \mathbb{R}^n : \|x - \hat{\delta}_{\hat{x}_{k-1}}\| \leq \delta_{\hat{x}_{k-1}} \},
\]

where \(\hat{d}_{k-1}\), \(\hat{\delta}_{\hat{x}_{k-1}}\), and \(\hat{x}_k\) are the hyperballs of compatible unknown inputs at time \(k - 1\), propagated, and updated states at time \(k\), correspondingly. In other words, we restrict the estimation errors to hyperballs of norm \(\delta\). In this setting, the observer design problem is equivalent to finding the centroids \(\hat{d}_{k-1}, \hat{\delta}_{\hat{x}_{k-1}}\), and \(\hat{x}_k\) as well as the radii \(\delta_{\hat{x}_{k-1}}\), \(\delta_{\hat{x}_{k-1}}\), and \(\delta_{\hat{x}_k}\) of the sets \(\hat{d}_{k-1}, \hat{\delta}_{\hat{x}_{k-1}}, \) and \(\hat{x}_k\), respectively. In addition, we limit our attention to observers for the centroids \(\hat{d}_{k-1}, \hat{\delta}_{\hat{x}_{k-1}}, \) and \(\hat{x}_k\) that belong to the class of three-step recursive filters given in References 9 and 11, with \(\hat{x}_{0|0} = \hat{x}_0\).

#### 4.1 | System transformation

To aid the observer design, we first carry out a similarity transformation based on singular value decomposition (SVD) to decompose the unknown input signal \(d_k\) of system (2) into two components \(d_{1,k}\) and \(d_{2,k}\), as well as to decouple the output equation in (2) into two components, \(z_{1,k}\) and \(z_{2,k}\), as follows.

Let \(p_{H} = \text{rk}(H)\). Using SVD, we rewrite the direct feedthrough matrix \(H\) as

\[
H = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix},
\]

where \(\Sigma \in \mathbb{R}^{p \times p} \) is a diagonal matrix of full rank, \(U_1 \in \mathbb{R}^{p \times p}\), \(U_2 \in \mathbb{R}^{(l-p) \times (l-p)}\), \(V_1 \in \mathbb{R}^{p \times p}\), and \(V_2 \in \mathbb{R}^{p \times (p - 2)}\), while \(U \equiv \begin{bmatrix} U_1 & U_2 \end{bmatrix}\) and \(V \equiv \begin{bmatrix} V_1 & V_2 \end{bmatrix}\) are unitary matrices. When there is no direct feedthrough, \(\Sigma, U_1\), and \(V_1\) are empty matrices, and \(U_2\) and \(V_2\) are...
and \( V_2 \) are arbitrary unitary matrices, while when \( p_H = p = l \), \( U_2 \) and \( V_2 \) are empty matrices, and \( U_1 \) and \( \Sigma \) are identity matrices.

Then, we decouple the unknown input into two orthogonal components:

\[
d_{1,k} = V_1^T d_k, \quad d_{2,k} = V_2^T d_k. \tag{9}
\]

Considering that \( V \) is unitary,

\[
d_k = V_1 d_{1,k} + V_2 d_{2,k}. \tag{10}
\]

and we can represent the system (2) as:

\[
x_{k+1} = f(x_k) + B_k u_k + G_1 d_{1,k} + G_2 d_{2,k} + W_k, \\
y_k = C x_k + D_k u_k + H_1 d_{1,k} + v_k, \tag{11}
\]

where \( G_1 \triangleq G V_1 \), \( G_2 \triangleq G V_2 \), and \( H_1 \triangleq H V_1 = U_1 \Sigma \). Next, the output \( y_k \) is decoupled using a nonsingular transformation \( T = \begin{bmatrix} T_1^T & T_2^T \end{bmatrix} = U_1^T \) to obtain \( z_{1,k} \in \mathbb{R}^{p_H} \) and \( z_{2,k} \in \mathbb{R}^{l-p_H} \) given by

\[
\begin{align*}
z_{1,k} & \triangleq T_1 y_k = U_1^T y_k \\
& = C_1 x_k + \Sigma d_{1,k} + D_{k,1} u_k + v_{1,k}, \\
z_{2,k} & \triangleq T_2 y_k = U_2^T y_k \\
& = C_2 x_k + D_{k,2} u_k + v_{2,k}. \tag{12}
\end{align*}
\]

where

\[
C_1 \triangleq U_1^T C, \quad C_2 \triangleq U_2^T C, \quad D_{k,1} \triangleq U_1^T D_k, \quad D_{k,2} \triangleq U_2^T D_k, \quad v_{1,k} \triangleq U_1^T v_k, \quad v_{2,k} \triangleq U_2^T v_k. \tag{13}
\]

This transformation is also chosen such that \( \left\| \begin{bmatrix} v_{1,k}^T & v_{2,k}^T \end{bmatrix} \right\| = \| U^T v_k \| = \| v_k \| \). As a result, we obtain the following transformed system (14), where the output (measurement) signal \( y_k \) is decoupled into two components \( z_{1,k} \) and \( z_{2,k} \), one with the full rank direct feedthrough matrix \( \Sigma \), and the other without direct feedthrough:

\[
x_{k+1} = f(x_k) + B_k u_k + G_1 d_{1,k} + G_2 d_{2,k} + W_k, \\
z_{1,k} = C_1 x_k + \Sigma d_{1,k} + D_{k,1} u_k + v_{1,k}, \\
z_{2,k} = C_2 x_k + D_{k,2} u_k + v_{2,k}. \tag{14}
\]

Remark 5. It is important to note that \( d_{2,k} \) cannot be estimated from \( y_k \) since it does not affect \( z_{1,k} \) and \( z_{2,k} \). Thus, in light of (14), we can only obtain a (one-step) delayed estimate of \( \hat{d}_{k-1} \). The reader may refer to Reference 10 for a more detailed discussion on when a delay is absent or when we can expect further delays.

### 4.2 Observer structure

Using the above transformation, we propose a three-step recursive observer structure to compute the state and input estimate sets, where we first estimate the unknown input signal (with one step lag), using the state estimates and the (decoupled) observations (cf. (15)–(17)), then we propagate the state estimates through the system dynamics to compute the \textit{a priori} state estimates (cf. (18) and (19)), and finally, we update the \textit{a posteriori} state estimates using the observations (cf. (20)). Our unknown input observer structure is chosen based on the decoupling of the unknown input signal \( d_k \) into two components, where the first component \( d_{1,k} \) can be “seen/observed” from the observations of the current time step and the other component \( d_{2,k} \) can be only “seen” with one time-step delay (cf. Remark 5). Then, we choose our observer gains \( M_1, M_2, L \) to cancel out the effects of the unknown inputs/attacks for these different components using \( M_1 \) and \( M_2 \), while \( L \) is used to minimize the \( H_{\infty} \) gain. The proposed three-step recursive observer structure is as follows:
Unknown input estimation (UIE):
\[
\hat{d}_{1,k} = M_1(z_{1,k} - C_1\hat{x}_{k|k} - D_{k,1}u_k),
\]
\[
\hat{d}_{2,k-1} = M_2(z_{2,k} - C_2\hat{x}_{k|k-1} - D_{k,2}u_k),
\]
\[
\hat{d}_{k-1} = V_1\hat{d}_{1,k-1} + V_2\hat{d}_{2,k-1}.
\]

Time update (TU):
\[
\hat{x}_{k|k-1} = f_{k-1}(\hat{x}_{k-1|k-1}) + B_{k-1}u_{k-1} + G_1\hat{d}_{1,k-1},
\]
\[
\hat{x}^*_{k|k} = \hat{x}_{k|k-1} + G_2\hat{d}_{2,k-1}.
\]

Measurement update (MU):
\[
\hat{x}_{k|k} = \hat{x}^*_{k|k} + L(y_k - C\hat{x}^*_{k|k} - D_ku_k)
\]
\[
= \hat{x}^*_{k|k} + L(z_{2,k} - C_2\hat{x}^*_{k|k} - D_{k,2}u_k).
\]

where \(L \in \mathbb{R}^{nxl}, \tilde{L} \equiv LU_1 \in \mathbb{R}^{nx(l-p_0)}, M_1 \in \mathbb{R}^{p_n \times p_U} \) and \(M_2 \in \mathbb{R}^{(p-p_n) \times (l-p_0)}\) are observer gain matrices that are designed according to Lemma 2 and Theorem 2 to minimize the “volumes” of the sets of compatible states and unknown inputs, quantified by the radii \(\delta_{k+1}^{\Pi}, \delta_{k}^{\Pi}\) and \(\delta_k^{\Pi}\). Note also that we applied \(L = LU_2U_2^T = LU_2^T\) from Lemma 2 into (20), where \(U_2\) is defined in (13). The resulting fixed-order set-valued observer (that will be further described in the following section) is summarized in Algorithm 1.

### 5  OBSERVER DESIGN AND ANALYSIS

In this section, we derive LMI conditions for designing observers that are quadratically stable in the absence of noise (Section 5.1) and optimal in the \(H_\infty\) sense in the presence of noise (Section 5.2) with uniformly bounded estimate radii (Section 5.3).

To design and analyze the observer, we first derive our observer error dynamics via the following Lemma 2. For conciseness, all proofs are provided in the Appendix.

**Lemma 2.** Consider system (2) (cf. Remark 2) and the observer (15)-(20). Suppose \(\text{rk}(C_2G_2) = p - p_U\), where \(C_2\) and \(G_2\) are given in (13). Then, designing observer matrix gains as \(M_1 = \Sigma^{-1}, \ M_2 = (C_2G_2)^\dagger, \ L_1 = 0\) and \(L = LU_2^T\) from Lemma 2 into (20), where \(U_2\) is defined in (13), yields \(M_1\Sigma = I\) and \(M_2C_2G_2 = I\), and leads to the following difference equation for the state estimation error dynamics (i.e., the dynamics of \(\hat{x}_{k|k} \triangleq x_k - \hat{x}_{k|k}\)):
\[
\hat{x}_{k+1|k+1} = (I - \tilde{L}C_2)\Phi(\Delta f_k - \Psi\hat{x}_{k|k}) + \mathcal{W}(\tilde{L})\overline{w}_k.
\]

where

\[
\Delta f_k \triangleq f_k(x_k) - f_k(\hat{x}_k), \quad \Phi \triangleq I - G_2M_2C_2,
\]
\[
\overline{w}_k \triangleq \left[ \begin{array}{c} \left( \frac{1}{\sqrt{2}} \right) v_k^T \\ \left( \frac{1}{\sqrt{2}} \right) v_{k+1}^T \end{array} \right]^T,
\]
\[
R \triangleq \begin{bmatrix} -\sqrt{2}\Phi G_1M_1T_1 & -\Phi W & -\sqrt{2}G_2M_2T_2 \\ 0_{(l-p_0) \times l} & 0_{(l-p_0) \times n} & -2T_2 \end{bmatrix},
\]
\[
Q \triangleq G_1M_1C_1, \quad \mathcal{W}(\tilde{L}) \triangleq (I - \tilde{L}C_2)R + \tilde{L}Q.
\]

Note that \(\overline{w}_k\) is chosen such that \(\lim_{k \to \infty} \frac{1}{k+1} \sum_{i=0}^{k} \overline{w}_i^T \overline{w}_i = \lim_{k \to \infty} \frac{1}{k+1} \sum_{i=0}^{k} \overline{w}_i^T \overline{w}_i + v_i^T v_i\). The result in (21) shows that we successfully decoupled/canceled out \(d_k\) from the error dynamics, otherwise there would be a potentially unbounded and unknown term in the error dynamics.
Algorithm 1. Simultaneous input and state observer (SISO)

1: Initialize:
\[ M_1 = \Sigma^{-1}; M_2 = (C_2G_2)^\dagger; \text{(cf. Lemma 2)} \]
\[ \Phi = I - G_2M_2C_2; \Psi = G_1M_1C_1; \]
Compute \( P, L, \rho \) via Theorem 2 and \( \theta_1, \theta_2, \bar{\eta}, \bar{\beta}, \bar{\alpha} \) via Theorem 4;
\[ \hat{x}_{0|0} = \hat{x}_0 = \text{centroid}(\hat{X}_0); \]
\[ \hat{d}_{1|0} = M_1(z_{1|0} - C_1\hat{x}_{0|0} - D_1\bar{\eta}_0); \]
\[ \delta_0 = \delta^* = \min \{||x - \hat{x}_{0|0}|| \leq \delta, \forall x \in \hat{X}_0\}; \]

2: for \( k = 1 \) to \( K \) do
\[ \triangleright \text{Estimation of } \hat{d}_{2,k-1} \text{ and } d_{k-1} \]
\[ \hat{x}_{k|k-1} = f_k(\hat{x}_{k-1|k-1}) + B_ku_{k-1} + G_k\hat{d}_{1,k-1}; \]
\[ \hat{d}_{2,k-1} = M_2(z_{2,k} - C_2\hat{x}_{k|k-1} - D_{k,2}u_k); \]
\[ \hat{d}_{1,k-1} = V_1\hat{d}_{1,k-1} + V_2\hat{d}_{2,k-1}; \]
\[ \delta_{k-1} = \beta\delta_{k-1} + \bar{\alpha}; \]
\[ \delta_{k-1} = \{d \in \mathbb{R}^l : ||d - \hat{d}_{k-1}|| \leq \delta_{k-1} \}; \]
\[ \triangleright \text{Time update} \]
\[ \hat{x}_{k|k} = \hat{x}_{k|k-1} + G_2\hat{d}_{2,k-1}; \]
\[ \triangleright \text{Measurement update} \]
\[ \hat{x}_{k|k} = \hat{x}_{k|k} + L(z_{2,k} - C_2\hat{x}_{k|k} - D_{k,2}u_k); \]
\[ \delta_{k,1} = \sqrt{(\delta_{0,1}^x)^2 \theta_{1}^k + \frac{\rho}{\lambda_{\text{min}}(P)}(\eta_{0}^w + \eta_{0}^v)} \sum_{i=1}^{k} \theta_{1}^{-1}; \]
\[ \delta_{k,2} = \delta_{0,2}^x \theta_{2}^k + \bar{\eta} \sum_{i=1}^{k} \theta_{2}^{-1}; \]
\[ \delta_{k} = \min(\delta_{k,1}, \delta_{k,2}); \]
\[ \hat{X}_k = \{x \in \mathbb{R}^n : ||x - \hat{x}_{k|k}|| \leq \delta_k \}; \]
\[ \triangleright \text{Estimation of } \hat{d}_{1,k} \]
\[ \hat{d}_{1,k} = M_1(z_{1,k} - C_1\hat{x}_{k|k} - D_{k,1}u_k); \]
end for

5.1 Stable observer design

In this section, we first investigate the existence of a stable observer in the form of (15)–(20) by providing necessary and sufficient conditions for quadratic stability of the observer for the system classes described in Section 2 by supposing for the moment that there is no exogenous bounded noise \( w_k \) and \( v_k \). Inspired by the definition of quadratic stability for nonlinear continuous-time systems in Reference 53, we formally define our considered notion of quadratic stability for nonlinear discrete-time systems.

Definition 8 (Quadratic stability). The nonlinear discrete-time dynamical system \( x_{k+1} = f_k(x_k) \), with the vector field \( f_k(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^n \), is quadratically stable, if it admits a quadratic positive definite Lyapunov function \( V_k = x_k^TPx_k > 0 \), with \( P \) being a positive definite matrix in \( \mathbb{R}^{nxn} \), such that the Lyapunov function increment \( \Delta V_k \triangleq V_{k+1} - V_k \) satisfies the following inequality for some \( \alpha \in [0, 1] \), for all \( k \in \mathbb{Z} \):

\[ \Delta V_k \leq -\alpha x_k^TPx_k. \] (22)

Remark 6. It can be shown that (22) implies that \( ||x_k|| \leq \frac{\lambda_{\text{max}}(P)}{\lambda_{\text{min}}(P)}(1 - \alpha)^k ||x_0|| \), which is exponentially decreasing with \( \alpha \in (0, 1) \), nonincreasing with \( \alpha = 0 \) or deadbeat with \( \alpha = 1 \). Note that there is also a slightly different notion of quadratic stability in the literature, for example, in References 54 and 55, with \( \Delta V_k \leq -\alpha x_k^Tx_k \), which implies \( ||x_k|| \leq \left( \frac{\lambda_{\text{max}}(P) - \alpha}{\lambda_{\text{min}}(P)} \right)^k ||x_0|| \), where the required condition on \( \alpha \) for stability is dependent on \( P \), making it slightly more complicated to perform a line search over \( \alpha \). Hence, in this article, we selected the notion of quadratic stability in (22), similar to Reference 53.

Now we are ready to state our first set of main results on necessary and sufficient conditions for the existence of quadratically stable observers for noiseless systems through the following theorem.
**Theorem 1** ( Necessary and sufficient conditions for quadratically stable observers). Consider system (2) (cf. Remark 2). Suppose there is no bounded noise \( w_k \) and \( v_k \) and all the conditions in Lemma 2 hold. Then, there exists a quadratically stable observer in the form of (15)–(20), if and only if there exist \( \alpha \in [0, 1] \), \( \kappa > 0 \) and matrices \( P, \Gamma, \tilde{Q}, Z \in \mathbb{R}^{nxn} \) and \( Y \in \mathbb{R}^{nx(l-p_H)} \) such that the following feasibility conditions hold:

\[
P > 0, \quad \Gamma > 0, \quad \tilde{Q} > 0,
\]
\[
\begin{bmatrix}
P & \tilde{Y}_1 \\
\tilde{Y}_1^T & M_1
\end{bmatrix} > 0,
\]
\[
\begin{bmatrix}
P & \tilde{Y}_2 \\
\tilde{Y}_2^T & M_2 + Z \Phi
\end{bmatrix} > 0,
\]
\[
\begin{bmatrix}
P & \tilde{Y}_1 \\
\tilde{Y}_1^T & M_3
\end{bmatrix} > 0,
\]
\[
\begin{bmatrix}
\Gamma & Z \\
Z^T & M_4 \Psi
\end{bmatrix} > 0,
\]

where \( \tilde{Y}_1 \triangleq (P - YC_2) \Phi, \tilde{Y}_2 \triangleq (P - YC_2) \Phi \Psi, \Phi, \Psi \) are defined in Lemma 2 and \( M_1, M_2, M_3, M_4 \) for Class 0 systems are given as follows:

0. If \( f_k(\cdot) \) is a Class 0 function with multiplier matrix \( M \triangleq \begin{bmatrix} M_{11} & M_{12} \\ M_{12}^T & M_{22} \end{bmatrix} \) and some \( \gamma > 0 \), then

\[
\begin{align*}
M_1 & \triangleq -\kappa M_{11} - \tilde{Q}, \\
M_2 & \triangleq -\kappa M_{22} + (1 - \alpha) P - \Gamma, \\
M_3 & \triangleq -\kappa M_{11}, \\
M_4 & \triangleq \kappa M_{12}.
\end{align*}
\]

Moreover, for system classes I, II–III, the \( M_{11}, M_{12}, \) and \( M_{22} \) matrices in (24) are given as follows:

(I) If \( f_k(\cdot) \) is a Class I function with Lipschitz constant \( L_f \), then

\[
M_{11} = -I, \quad M_{12} = 0, \quad M_{22} = L_f^2 I.
\]

(II) If \( f_k(\cdot) \) is a Class II function with multiplier matrix \( A \), then

\[
M_{11} = -I, \quad M_{12} = A, \quad M_{22} = -A^T A.
\]

(III) If \( f_k(\cdot) \) is a Class III function with constituent matrices \( A_i \in \mathbb{R}^{dxn}, \forall i \in \{1, \ldots, N\} \) and \( \sigma_m \triangleq \max_{i \in 1 \ldots N} \| A_i \|, \) then

\[
M_{11} = -I, \quad M_{12} = 0, \quad M_{22} = \sigma_m^2 I.
\]

Furthermore, no quadratically stable estimator can be designed if \( \gamma < 0 \).

**Remark 7.** The feasibility conditions in Theorem 1 can be easily verified by applying line search/bisection over \( \alpha \) and solving the corresponding LMIs for \( \kappa, P, \Gamma, \tilde{Q}, \tilde{Z}, \) and \( Y \), given \( \alpha \).

**Remark 8.** Note that although the computational complexity of the SDP in (23) grows when the state space dimension increases, this is usually less of a concern since the LMI-based SDPs are solved “offline.” More precisely, to design a stable observer, the SDP in (23) that has eight LMI constraints and \( 3(N^2 - N + n) + n^2 + n(l - p_H) + 2 \) real decision variables needs to be only solved “once.”

Theorem 1 provides powerful tools in terms of necessary and sufficient conditions for designing quadratically stable observers. When the LMIs in (23) do not hold, it equivalently implies that there does not exist any quadratically stable observer for that particular system. However, in such cases, one may still be able to design a Lyapunov stable observer, given the fact that quadratic and Lyapunov stability are not equivalent for general nonlinear systems (since Lyapunov stability, in its most general sense, hinges upon admitting any form of Lyapunov functions and not necessarily a quadratic form). This motivates us to derive necessary conditions in terms of LMI “infeasibility” conditions for Lyapunov stability.
of the observer. If these necessary conditions are feasible, then we know for certain that no stable observer, in the most general sense of stability, can be designed.

**Proposition 10** (Necessary condition for observer Lyapunov stability). Consider system (2) (cf. Remark 2) and the observer (15)–(20). Suppose there is no bounded noise \( w_k \) and \( v_k \) and all the conditions in Lemma 2 hold. Then, the observer error dynamics is Lyapunov stable, only if the following LMIs are always infeasible for all \( 0 < \bar{P} \in \mathbb{R}^{n \times n}, Y \in \mathbb{R}^{n \times (l-p_H)} \), \( 0 < \bar{\gamma} \in \mathbb{R}^{(l-p_H) \times (l-p_H)} \) and \( 0 < \bar{\eta} < 1 \).

\[
\begin{bmatrix}
I - \bar{\gamma} & 0 & 0 \\
0 & \bar{\gamma} & \bar{Y}^T \\
0 & \bar{Y} & \bar{P}
\end{bmatrix} \succ 0, \quad \begin{bmatrix}
\bar{\Pi}_{11} & \bar{\Pi}_{12} \\
\bar{\Pi}_{12}^T & \bar{\Pi}_{22}
\end{bmatrix} \succ 0,
\]

(28)

where \( \bar{S} \triangleq \bar{P} - C_2^T \bar{Y}^T - \bar{Y}C_2 \) and \( \bar{\Pi}_{11}, \bar{\Pi}_{12}, \) and \( \bar{\Pi}_{22} \) for Class 0 systems are defined as follows:

0. If \( f_k(\cdot) \) is a Class 0 function with multiplier matrix \( M \triangleq \begin{bmatrix} M_{11} & M_{12} \\ M_{12}^T & M_{22} \end{bmatrix} \) and some \( \gamma \geq 0 \), then

\[
\begin{align*}
\bar{\Pi}_{11} & \triangleq \Phi^T(\bar{S} - (1 - \bar{\eta})C_2^T C_2)\Phi - M_{11}, \\
\bar{\Pi}_{12} & \triangleq -\Phi^T\bar{S}\Phi - M_{12}, \\
\bar{\Pi}_{22} & \triangleq \Psi^T \Phi^T(\bar{S} - (1 - \bar{\eta})C_2^T C_2)\Phi \Psi - \bar{P} - M_{22}.
\end{align*}
\]

Moreover, for system classes I, II–III, the \( M_{11}, M_{12}, \) and \( M_{22} \) matrices in (29) are given by (25), (26), and (27), respectively.

It is worth mentioning that if \( f_k(\cdot) \) is a Class III function, then we can provide sufficient conditions for the existence of Lyapunov stable observers as well as necessary conditions that are conveniently testable. The latter are beneficial in the sense that if they are not satisfied, the designer knows a priori that there does not exist any stable observer for those LPV systems with unknown inputs/attacks. The conditions are formally derived in the following Lemma 3.

**Lemma 3.** Suppose \( f_k(\cdot) \) is a Class III function and all the conditions in Lemma 2 hold. Then, there exists a stable observer for the system (3), with any sequence \( \{\lambda_{i,k}\}_{k=0}^\infty \) for all \( i \in \{1, 2, \ldots, N\} \) that satisfies \( 0 \leq \lambda_{i,k} \leq 1, \sum_{i=1}^N \lambda_{i,k} = 1, \forall k \), if \( (\bar{A}_k, C_2) \) be uniformly detectable\(^\dagger\) for each \( k \), and only if all constituent LTI systems \( (A^i, G, C, H) \), \( \forall i \in \{1 \ldots N\} \), are strongly detectable (cf. Definition 1), where \( \bar{A}_k \triangleq \Phi \sum_{i=1}^N \lambda_{i,k} A^i \Phi^{-1} \), with \( \Phi \) and \( \Psi \) defined in Lemma 2.

**Corollary 3.** There exists a stable simultaneous/joint state and input set-valued observer for the LTI system (7), through (15)–(20), if and only if the tuple \( (A, G, C, H) \) is strongly detectable and only if \( \text{rk}(C_2 G_2) = p - p_H \). Moreover, the observer gain matrices can be designed as \( M_1 = \Sigma^{-1} \), \( M_2 = (C_2 G_2)^\dagger \) and \( L = LU_2^T \) and \( L = P^{-1}Y \), where \( P > 0 \) and \( Y \) solve the following feasibility program with LMI constraints:

\[
\begin{align*}
\text{Find} \ (P > 0, Y ) \\
\text{s.t.} \quad \begin{bmatrix}
P & \Lambda \\ \Lambda^T & P
\end{bmatrix} \succeq 0,
\end{align*}
\]

with \( \Lambda \triangleq (A - \Psi)^T (P - C_2^T Y^T) \) and \( \Phi \) and \( \Psi \) defined in Lemma 2.

### 5.2 \( H_\infty \) observer design

The goal of this section is to provide additional sufficient conditions to guarantee optimality of the observers in the \( H_\infty \) sense in the presence of exogenous noise. We first define our considered notion of optimality via the following Definition 9.

**Definition 9** (\( H_\infty \) observer). Let \( T_{\tilde{x},w,v} \) denote the transfer function matrix that maps the noise signals \( \tilde{w}_k \triangleq [w_k^T \ v_k^T]^T \) to the updated state estimation error \( \tilde{x}_k \triangleq x_k - \tilde{x}_k \). For a given noise attenuation level \( \rho \in \mathbb{R}_+ \), the observer
performance satisfies $\mathcal{H}_\infty$ norm bounded by $\rho$, if $\|T_{x,w,\nu}\|_\infty \leq \rho$, that is, the maximum average signal power amplification is upper-bounded by $\rho^2$:

$$\lim_{k \to \infty} \frac{1}{k+1} \sum_{i=0}^{k} \|x_i^T x_i\| \leq \|T_{x,w,\nu}\|_\infty^2 \leq \rho^2.$$  \hfill (30)

Now, we present our second set of main results on designing stable and optimal observers in the minimum $\mathcal{H}_\infty$ sense. In particular, starting from the observer error dynamics obtained in (21), we define the quadratic candidate Lyapunov function $V_R^n \triangleq x_k^T P x_k$ and show that its increment, \( \Delta V_R^n \equiv V_{R,n+1} - V_R^n = \Delta V^{aw} + \Delta r_k \), can be rewritten as a summation of the increment of the noise-less Lyapunov function $V_R^{aw}$, used for the stability analysis in the proof of Theorem 1, and the error term \( \Delta r_k \equiv 2(\Delta f_k^T x_k)(\Phi \bar{P}(I - \hat{L}C_k)^T PW(L)\bar{w}_k + \bar{w}_k^T W(L)^T PY(L)\bar{w}_k) \). Then, leveraging S-procedure, Schur complement and several similarity transformations (cf. proof of Theorem 2 in Appendix A.14 for more details), we derive LMI conditions for each system class that imply

$$\Delta r_k \leq \Delta r_k - \rho^2 \bar{w}_k^T \bar{w}_k + x_k^T (I - aP)\bar{x}_k \leq 0.$$  \hfill (31)

Combining these two results and the fact that $\Delta V^{aw} \leq -a x_k^T P \bar{x}_k$ (follows from Theorem 1, that is, the LMI conditions in (23)–(27)), we obtain

$$\Delta V^n_R \leq \rho^2 \bar{w}_k^T \bar{w}_k - x_k^T x_k.$$  \hfill (34)

Summing up both sides of the above inequality from zero to infinity, returns $V_R^n - V_R^n_0 \leq \rho^2 \sum_{k=0}^{\infty} \bar{w}_k^T \bar{w}_k - \sum_{k=0}^{\infty} x_k^T x_k = \rho^2 \sum_{k=0}^{\infty} \bar{w}_k^T \bar{w}_k - \sum_{k=0}^{\infty} x_k^T x_k$, where at each time step $k$, $\bar{w}_k^T = [w_k^T, v_k^T]^T$. Then, it follows from setting the initial conditions to zero that $\sum_{k=0}^{\infty} x_k^T x_k \leq \rho^2 \sum_{k=0}^{\infty} \bar{w}_k^T \bar{w}_k$. Hence, by solving an SDP, where we minimize $\rho$, subject to (32), (33), which is a combination of the LMI conditions to guarantee (31) and the LMIs in (23)–(27) that are required to guarantee stability, we design stable and optimal observers in the minimum $\mathcal{H}_\infty$ sense, that is, we minimize $\rho$, the norm of the transfer function in (30). Interested readers are referred to the proof of Theorems 1 and 2 in Appendices A.11 and A.14, respectively, for more details.

**Theorem 2** ($\mathcal{H}_\infty$ observer design). Consider system (2) (cf. Remark 2), the observer (15)–(20) and a given $\rho > 0$. Suppose all the conditions in Theorem 1 hold and let $\Phi, \Psi, Q$ and $R$ be defined as in Lemma 2 and $\Omega \triangleq C_R - Q$. Then, with the gain $\bar{L} = P^{-1}Y$, we obtain a quadratically stable observer with $\mathcal{H}_\infty$ norm bounded by $\rho$, if the LMIs in (23) hold with some $P > 0$ and $Y = \bar{P}\bar{L}$, and there exist $0 \leq \Gamma \in \mathbb{R}^{nxn}$ and $\epsilon_1, \epsilon_2 > 0$ such that:

$$\Pi \triangleq \begin{bmatrix} I - \Gamma & 0 & 0 \\ 0 & P & Y \\ 0 & Y^T & I \end{bmatrix} \succeq 0, \quad \mathcal{N} \triangleq \begin{bmatrix} \mathcal{N}_{11} & * & * \\ \mathcal{N}_{21} & \mathcal{N}_{22} & * \\ \mathcal{N}_{31} & \mathcal{N}_{32} & \mathcal{N}_{33} \end{bmatrix} \succeq 0,$$  \hfill (32)

where

$$\begin{align*}
\mathcal{N}_{11} & \triangleq \rho^2 I + 2R^T \Gamma \Omega - R^T PR - \Omega^T (\Gamma + (\epsilon_1^{-1} + \epsilon_2^{-1})I)\Omega, \\
\mathcal{N}_{21} & \triangleq \Psi^T (PR - Y\Omega - C'_2 Y^TR), \\
\mathcal{N}_{31} & \triangleq \Phi^T (Y\Omega + C'_2 Y^TR - PR),
\end{align*}$$  \hfill (33)

and $\mathcal{N}_{22}, \mathcal{N}_{32}$ and $\mathcal{N}_{33}$ are defined for Class 0 systems as follows:

0. If $f(\cdot)$ is a Class 0 function with multiplier matrix $\mathcal{M} \triangleq \begin{bmatrix} M_{11} & M_{12} \\ M_{12}^T & M_{22} \end{bmatrix}$ and some $\gamma \geq 0$, then

$$\begin{align*}
\mathcal{N}_{22} & \triangleq -I + aP - \epsilon_1 \Psi^T C'_2 C_2 \Phi \Psi - M_{22}, \\
\mathcal{N}_{32} & \triangleq -M_{12}, \\
\mathcal{N}_{33} & \triangleq -\epsilon_2 \Phi^T C'_2 C_2 \Phi - M_{11}.
\end{align*}$$  \hfill (34)
Moreover, for system classes I, II–III, the $M_{11}$, $M_{12}$ and $M_{22}$ matrices in (34) are given by (25), (26) and (27), respectively. Finally, the minimum $H_\infty$ bound can be found by solving the following semi-definite program (SDP) (with line searches over $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$):

$$
(\rho^*)^2 = \min_{P > 0, \Gamma > 0, y, \rho' > 0, 0 \leq a \leq 1, x > 0} \rho^2
$$

s.t. (23), (32) hold, \hfill (35)

where $\rho^2$ is a decision variable. If this SDP is feasible, then the infinity norm of the transfer function matrix $T_{x,w,v}$ satisfies $\|T_{x,w,v}\|_\infty \leq \rho^*$. This bound is obtained by applying the observer gain $\hat{L}^* = P^{-1}Y^*$, where $(P^*, Y^*, \Gamma^*)$ solves the above SDP.

Remark 9. Similar to Remark 8, the computational burden imposed by the SDP in (35) with 13 LMI constraints and $4(n^2-n/2 + n) + n^2 + n(I - p_{ll}) + 4$ real decision variables is not a huge concern because to design an $H_\infty$ observer, we only need to solve an SDP once “offline” to obtain the observer gain $\hat{L}$ for the observer in (15)–(20) (as well as $M_1 = \Sigma^{-1}$; $M_2 = (C_2G_2)^T$ from Lemma 2).

5.3 Radii of estimates and convergence of errors

Although the initial state and noise/disturbance are assumed to be bounded, the estimates radii can still potentially become unbounded due to the unbounded attacks and nonlinear dynamics. It is similar to deterministic systems (including noise-less linear systems) where the estimation errors can grow unbounded if the observer is not stable. Hence, the main purpose of this section is to guarantee boundedness (or “stability”) of the radii and also to provide closed-form expressions for upper bounds/over-approximations of the estimate radii and (iii) finding sufficient conditions for the convergence of the upper bound sequences, as well as their steady-state values (if they exist).

It is worth mentioning that for linear time-invariant systems, strong detectability of the system (cf. Definition 1) is a sufficient condition for the convergence of the radii $\delta^s$ and $\delta^d$ to steady state, but it is less clear for general nonlinear systems. Notice that if $f_k(\cdot)$ is a Class III function, that is, in the LPV case, even strong detectability of all constituent LTI systems does not guarantee that the radii converge. The reason is that the convergence hinges on the stability of the product of time-varying matrices (cf. proof of Theorem 4), which is not guaranteed even if all the multiplicands are stable.

To address the existence of uniformly bounded radii for the proposed observer designs for the nonlinear systems we consider, we first define the notion of uniformly bounded-input bounded-state systems.

Definition 10 (UBIBS systems\textsuperscript{section 3.2}). A dynamic system is uniformly bounded-input bounded-state (UBIBS), if bounded initial states $x_0$ and bounded (disturbance/noise) inputs $u$ produce uniformly bounded trajectories, that is, there exist two $K$-functions $\sigma_1$ and $\sigma_2$ such that

$$
\sup_k ||x(k, x_0, u)|| \leq \max \{ \sigma_1(||x_0||), \sigma_2(||u||) \}.
$$

Now, we are ready to state our results on the uniform boundedness of the estimate radii.

Theorem 3 (Uniformly bounded estimate radii). Consider system (2) (cf. Remark 2) and the observer (15)–(20). Suppose all the conditions in Theorem 2 hold. Then, the state estimation radii/error dynamics (21) is a UBIBS system with noise as exogenous inputs. In other words, bounded initial state errors and noise produce uniformly bounded trajectories of errors, that is, there exist $K$-functions $\sigma_1$ and $\sigma_2$ such that

$$
\sup_k ||\tilde{x}(k)|| \leq \max \left\{ \sigma_1(||\tilde{x}(k)||), \sigma_2 \left( \sqrt{\eta^2 + \eta^2} \right) \right\}.
$$

Moreover, (21) admits a $K$-asymptotic gain, that is, there exist $K$-function $\gamma_a$ such that

$$
\lim sup_{k \to \infty} ||\tilde{x}(k)|| \leq \gamma_a \left( \lim sup_{k \to \infty} ||\tilde{w}(k)|| \right).
$$

where $\lim sup_{n \to \infty} x_n = \lim_{n \to \infty}(\sup_{m \geq n} x_m)$ denotes the limit superior of the sequence $\{x_n\}_{n=1}^\infty$. 


The above Theorem 3 guarantees uniform boundedness of the estimate radii, if an $H_\infty$ observer in the form of (15)–(20) exists and can be designed through Theorem 2. Next, we are interested in deriving closed-form expressions for the upper bound/over-approximation of the uniformly bounded estimate radii, that is, upper bound sequences for the resulting sequences of radii $\{\delta_k^\alpha\}_{k=1}^\infty$ and $\{\delta_k^d\}_{k=1}^\infty$, when using our proposed observer for the different classes of systems. We also discuss some sufficient conditions for the convergence of the over-approximations of the estimate radii to steady state.

**Theorem 4** (Upper bounds of the radii of estimates). Consider system (2) (cf. Remark 2) along with the observer (15)–(20). Suppose the conditions of Theorem 2 hold. Let $\mathcal{R} \triangleq - (\Psi \Phi G_1 M_1 T_1 + \Psi G_2 M_2 T_2 + LT_2)$, $\overline{\alpha} \triangleq \| V_2 M_2 C_2 \| \eta_w + \|(V_2 M_2 C_2 G_1 - V_1) M_1 T_1\| + \| V_2 M_2 T_2 \| \eta_w$ and $\widetilde{\eta} \triangleq \| \mathcal{R} \| \eta_w + \| \Psi \Phi \| \eta_w$ with $\Phi$ and $\Psi$ defined in Lemma 2 and $T_1, T_2$ given in Section 4.1. Then, the upper bound sequences for the estimate radii, denoted by $\delta_k^\alpha$ and $\delta_k^d$, can be obtained as:

$\delta_k^\alpha \leq \min(\overline{\delta}_{k,1}^\alpha, \overline{\delta}_{k,2}^\alpha), \quad (36)$

$\delta_k^d \leq \overline{\delta}_{k-1}^d \triangleq \beta \overline{\delta}_{k-1} + \overline{\alpha}, \quad (37)$

where

$\overline{\delta}_{k,1}^\alpha \triangleq \sqrt{(\delta_0^\alpha)^2 \theta_0^1 + \frac{\beta^2}{\lambda_{\min}(P)} (\eta_0^2 + \eta_1^2) \sum_{i=1}^k \theta_i^{1-1}}, \quad (38)$

$\theta_1 \triangleq \frac{|\lambda_{\max}(P) - 1|}{\lambda_{\min}(P)}, \quad (39)$

$\overline{\delta}_{k,2}^\alpha \triangleq \delta_0^\alpha \theta_0^2 + \eta \sum_{i=1}^k \theta_i^{2-1}, \quad (40)$

and $\theta_2$ and $\beta$ are defined for the different function classes as follows:

(I) If $f(\cdot)$ is a Class I function, then

$\theta_2 \triangleq (L_f + \| \Psi \|)(I - L C_2) \Phi, \quad (41)$

$\beta \triangleq \| V_1 M_1 C_1 - V_2 M_2 C_2 \Phi \| + L_f \| V_2 M_2 C_2 \|.$

(II) If $f(\cdot)$ is a Class II function, then

$\theta_2 \triangleq (\lambda_{\max}(A^T A) + \| \Psi \|)(I - L C_2) \Phi, \quad (42)$

$\beta \triangleq \| V_1 M_1 C_1 - V_2 M_2 C_2 \Phi \| + \lambda_{\max}(A^T A) \| V_2 M_2 C_2 \|.$

(III) If $f(\cdot)$ is a Class III function, then

$\theta_2 \triangleq \max_{i \in \{1, 2, \ldots, N\}} \| A_{c,i} \|, \quad (43)$

$\beta \triangleq \max_{i \in \{1, 2, \ldots, N\}} \| V_1 M_1 C_1 + V_2 M_2 C_2 A_{c,i} \|.$

with $A_{c,i} \triangleq (I - L C_2) \Phi (A^i - \Psi)$, for all $i \in \{1 \ldots N\}$ and $V_1, V_2$ given in Section 4.1.

Furthermore, the upper bound sequences for the estimate radii are convergent if $\theta \triangleq \min(\theta_1, \theta_2) < 1$ (equivalently, if $\theta_1 < 1$ or $\theta_2 < 1$), and in this case, the steady-state upper bounds of the estimate radii are given by:

$\lim_{k \to \infty} \delta_k^\alpha = \delta_\infty^\alpha \triangleq \begin{cases} \overline{\delta}_{\infty,1}^\alpha, & \text{if } \theta_1 < 1, \theta_2 \geq 1, \\ \overline{\delta}_{\infty,2}^\alpha, & \text{if } \theta_1 \geq 1, \theta_2 < 1, \\ \min(\overline{\delta}_{\infty,1}^\alpha, \overline{\delta}_{\infty,2}^\alpha), & \text{if } \theta_1 < 1, \theta_2 < 1, \end{cases}$

$\lim_{k \to \infty} \delta_k^d = \beta \delta_\infty^\alpha + \overline{\alpha}, \quad (44)$
where $\tilde{\delta}_{x,1} = \rho \sqrt{\frac{n_2^2 + n_1^2}{\delta_{m}(P(1 - \theta_1))}}$ and $\tilde{\delta}_{x,2} \triangleq \frac{\pi}{1 - \theta_1}$.

**Corollary 4.** If $f_k(\cdot)$ is a Class III function and the conditions of Theorem 2 hold, then, the upper bound sequences for the estimate radii, i.e., $\tilde{\delta}_k^x$ and $\tilde{\delta}_{k-1}^d$, computed in (36)–(40), are convergent if $\|A_{e,d}\| < 1$ for all $i \in \{1, 2, \ldots, N\}$, where $A_{e,d} \triangleq (I - LC_I)\Phi(A^I - \Psi)$, with $\Phi$ and $\Psi$ defined in Lemma 2.

**Remark 10.** According to Theorem 3, if the necessary and sufficient conditions in Theorem 2 hold, that is, when the observer is quadratically stable and optimal in the sense of $H_{\infty}$, the sequences of estimate radii, $\{\tilde{\delta}_k^x, \tilde{\delta}_{k-1}^d\}_{k=1}^{\infty}$, are uniformly bounded, regardless of the value of $\theta_1$ or $\theta_2$. Consequently, the sequences of errors, $\{\tilde{x}_k, \tilde{\delta}_{k-1}\}_{k=1}^{\infty}$, are also uniformly bounded and do not diverge. On the other hand, the closed-form (potentially conservative) upper bound sequences, $\{\tilde{\delta}_k, \tilde{\delta}_{k-1}\}_{k=1}^{\infty}$, may diverge even when $\{\tilde{x}_k, \tilde{\delta}_{k-1}\}_{k=1}^{\infty}$ are uniformly bounded, and a sufficient condition for their convergence is that $\theta = \min(\theta_1, \theta_2) < 1$, that is, $\theta_1 < 1$ and/or $\theta_2 < 0$.

We conclude this section by stating a proposition with which we can tradeoff between observer optimality (i.e., the noise attenuation level) and convergence of the upper bound sequences for the error radii by adding some additional LMI s to the conditions in (23) and (32), and solving the corresponding mixed-integer SDP.

**Proposition 11 (Convergence of upper bound sequences).** Consider system (2) (cf. Remark 2) and suppose that the assumptions in Lemma 2 hold. Then, solving the following mixed-integer SDP:

$$
(\rho^{**})^2 = \min_{\{P > 0, \Gamma > 0, \gamma, \theta > 0, \omega < 1, \varepsilon_1 > 0, \varepsilon_2, > 0, \varepsilon_3 > 0, \varepsilon_4 > 0, \varepsilon_5 > 0\}} \rho^2
$$

s.t. (23), (32) hold,

$$
\kappa_1 I \preceq P \preceq \kappa_2 I,
$$

$$
(\kappa_1 \geq 1, \kappa_2 - \kappa_1 < 1) \lor (\kappa_2 \leq 1, \kappa_1 > 0.5),
$$

guarantees that $\theta_1 < 1$ and thus, $\theta < 1$, where $\theta_1$ and $\theta$ are given in Theorem 4, and results in a quadratically stable observer in the form of (15)–(20), with convergent upper bound sequences for the radii $\{\tilde{\delta}_k^x, \tilde{\delta}_{k-1}^d\}_{k=1}^{\infty}$ in the form of $\tilde{\delta}_k^x$ and $\tilde{\delta}_{k-1}^d$ in (36) and (37), respectively, with noise attenuation level $\rho^{**}$ when using the observer gain $\hat{L}^{**} = (P^{**})^{-1}Y^{**}$, where $(P^{**}, Y^{**})$ are solutions to the above mixed-integer SDP.

Note that the above mixed-integer SDP can also be solved using two independent SDPs with each the disjunctive constraints (denoted with $\lor$) and selecting the solution corresponding to the smaller $\rho^{**}$. Moreover, it is worth mentioning that although the designed observer may not be optimum in the minimum $H_{\infty}$ sense when using the mixed-integer SDP in Proposition 11, we can instead guarantee the steady-state convergence of the closed-form upper bound sequences of the estimate radii.

### 6 | SIMULATION RESULTS AND COMPARISON WITH BENCHMARK OBSERVERS

Two simulation examples are considered in this section to demonstrate the performance of the proposed observer. In the first example, where the dynamic system belongs to Classes I and II, we consider simultaneous/joint input and state estimation problem and design observers for each class to study their performances. Our second example is a benchmark dynamical Lipschitz continuous (i.e., Class I) system, where we compare the results of our observer with two other existing observers in the literature.\textsuperscript{30,32} We consider two different scenarios, one with a bounded unknown input, and the other with an unbounded unknown input. The results show that in the unbounded input scenario, when applying the observers in References 30 and 32, the estimation errors diverge, while as expected from our theoretical results, the estimation errors of our proposed observer converge to steady state values. Finally, for the sake of completeness, we include the simulation results from our preliminary work,\textsuperscript{36} where we applied the proposed observer to an LPV, that is, Class III system.

### 6.1 | Single-link flexible-joint robotic system

We consider a single-link manipulator with flexible joints,\textsuperscript{13,59} where the system has four states. We slightly modify the dynamical system described in Reference 13, by ignoring the dynamics for the unknown inputs...
Comparison with benchmark observers

Of (3) with \( P \) that our set-valued estimates still converge, that is, our observer remains stable for all 50 randomly chosen initial values, both (point) state and unknown input estimates, while the observer in Reference 32, only obtains (point) state estimates. II. Further, we observe from Figure 2 that our proposed \([0, 1, 2, 1]^T\), \( C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \ W = I, \ D = 0_{2 \times 1}, \ T_s = 0.01, \ H = T_s \begin{bmatrix} 1.1 & 2 \end{bmatrix}^T, \) and \( \eta_w = \eta_\nu = 0.1. \) The unknown input signal is depicted in Figure 2. Vector field \( f(\cdot) \) is a Class I function with \( L_f = 3.33 T_s \| \text{diag}(0, 0, 0, 0) \| = 3.33 T_s \) (cf. Reference 13), as well as a Class II function, with \( A = A \) and \( \gamma = 0 \) (cf. Lemma 1). It is also a Class 0 function with \( \mathcal{M} = \begin{bmatrix} I & 0 \\ 0 & (3.33 T_s)^2 I \end{bmatrix} \geq \begin{bmatrix} I & 0 \\ 0 & (3.33 T_s)^2 I \end{bmatrix}, \) by Propositions 2–5. Solving the SDPs in Theorem 2 corresponding to system classes 0, I–II return \( P^*_0 = \begin{bmatrix} 1.4458 & -1.5232 & -0.3419 & -0.2265 \\ -1.5232 & 2.5753 & -0.2546 & -0.1828 \\ -0.3419 & -0.2546 & 1.2159 & -0.1475 \\ -0.2265 & -0.1828 & -0.1475 & 1.2605 \end{bmatrix}^T \) for \( Y^*_0 = \begin{bmatrix} 0.4348 & 0.3372 & 0.1376 & 0.1243 \end{bmatrix}^T, \ a^*_0 = 0.750, \ \rho^*_0 = 1.1781, \) and \( \tilde{L}_0^* = \begin{bmatrix} 1.2471 & 0.8705 & 0.4783 & 0.2947 \end{bmatrix}^T \) for Class 0, \( P^*_I = \begin{bmatrix} 1.6684 & -1.8242 & -0.4606 & -0.2268 \\ -1.8242 & 2.8284 & -0.2860 & -0.0424 \\ -0.4606 & -0.2860 & 1.2086 & -0.0628 \\ -0.2208 & -0.0424 & -0.0628 & 1.2088 \end{bmatrix} \) \( Y^*_I = \begin{bmatrix} 0.6464 & 0.4422 & 0.0420 & 0.0264 \end{bmatrix}^T, \ a^*_I = 0.825, \ \rho^*_I = 0.9436, \) and \( \tilde{L}_I^* = \begin{bmatrix} 1.2620 & 0.4288 & 0.4244 & 0.2667 \end{bmatrix}^T \) for Class I, \( P^*_II = \begin{bmatrix} 2.2823 & -2.7731 & -1.0065 & -0.5036 \\ -2.7731 & 4.7225 & -0.1605 & -0.9806 \\ -1.0065 & -0.1605 & 4.5760 & -0.3148 \\ -0.5036 & -0.9806 & -0.3148 & 4.0000 \end{bmatrix}^T, \ Y^*_II = \begin{bmatrix} -0.9692 & 1.0644 & 0.0259 & 0.0661 \end{bmatrix}^T, \ a^*_II = 0.793, \ \rho^*_II = 0.9783, \) and \( \tilde{L}_II^* = \begin{bmatrix} 0.4605 & 0.1837 & 0.9321 & 0.3519 \end{bmatrix}^T \) for Class II. Further, we observe from Figure 2 that our proposed \( \mathcal{H}_\infty \) observer, that is, Algorithm 1, is able to find set-valued estimates of the states and unknown inputs, for \((\mathcal{M}, \gamma, \cdot)\)-QC (Class 0), Lipschitz continuous (Class I) and \((\mathcal{M}, \gamma, \cdot)\)-QC* (Class II) functions. The actual estimation errors are also within the predicted upper bounds (cf. Figure 3), which converge to steady-state values as established in Theorem 4. Furthermore, Figures 2 and 3 show that in this specific example system, estimation errors and their radii are tighter when applying the obtained observer gains for Class I (i.e., Lipschitz) functions, when compared to applying the ones corresponding to the Class 0 (i.e., \((\mathcal{M}, \gamma, \cdot)\)-QC) and Class II (i.e., \((\mathcal{M}, \gamma, \cdot)\)-QC*) functions.

6.2 Comparison with benchmark observers

In this section, we illustrate the effectiveness of our simultaneous input and state set-valued observer (SISO), by comparing its performance with two benchmark observers in References 30 and 32. The designed estimator in Reference 30 calculates both (point) state and unknown input estimates, while the observer in Reference 32, only obtains (point) state estimates. For comparison, we apply all the three observers on a benchmark dynamical system in Reference 30, which is in the form of (3) with \( n = 2, m = l = p = 1, f(x) = \begin{bmatrix} -0.42 x_1 + x_2 \\ -0.6 x_1 - 1.25 \tanh(x_1) \end{bmatrix}, \ G = \begin{bmatrix} 1 & \ -0.65 \end{bmatrix}^T, \ B = D = H = 0_{1 \times 1}, \ C = \begin{bmatrix} 0 & 1 \end{bmatrix}, \ W = I, \ \eta_w = 0.2 \) and \( \eta_\nu = 0.1. \) The vector field \( f(\cdot) \) is Lipschitz continuous (i.e., Class I) with \( L_f = 1.1171. \) We consider two scenarios for the unknown input. In the first, we consider a random signal with bounded norm, that is, \( \| \delta_i \| \leq 0.2 \) for the unknown input \( \delta_i, \) while \( \delta_i \) in the second scenario is a time-varying signal that becomes unbounded as time increases. As is demonstrated in Figures 4 and 5, in the first scenario with bounded unknown inputs, the set estimates of our approach (i.e., SISO estimates) converge to steady-state values and the point estimates of the two benchmark approaches\(^{30,32} \) are within the predicted upper bounds and exhibit a convergent behavior for all 50 randomly chosen initial values (cf. Figure 5). In this scenario, the two benchmark approaches result in slightly better performance than SISO, since they benefit from the additional assumption of bounded input.

More interestingly, considering the second scenario with unbounded unknown inputs, Figures 6 and 7 demonstrate that our set-valued estimates still converge, that is, our observer remains stable for all 50 randomly chosen initial values, with \( P^* = \begin{bmatrix} 1.8086 & 1.4022 \\ 1.4022 & 5.2068 \end{bmatrix}, \ Y^* = \begin{bmatrix} -0.2282 & 0.6664 \end{bmatrix}^T, \ L^* = \begin{bmatrix} -0.2604 & 0.2064 \end{bmatrix}, \ a^* = 0.8875, \) and \( \rho^* = 1.7336, \) while
**Figure 2** Actual states $x_1, x_2, x_3, x_4$ and input $d$, as well as their Class 0, Class I, and Class II estimates (i.e., the obtained estimates by applying the corresponding gains for Classes 0–II in Theorem 2)

**Figure 3** Estimation errors and their upper bounds for Class 0 ($\mathcal{M}, \gamma$-QC), Class I (Lipschitz) and Class II ($\mathcal{A}, \gamma$-QC*) functions
the estimates of the two benchmark approaches exceed the boundaries of the compatible sets of states and inputs after some time steps of our approach and display a divergent behavior for all initial values (cf. Figure 7).

6.3 A Class III (LPV) system

Finally, for the sake of completeness, we also present the simulation results from our preliminary work,\textsuperscript{36} where we applied the proposed observer to an LPV system (i.e., Class III system; cf. Definition 3) that consists of a convex combination of two constituent linear time-invariant strongly detectable subsystems in the form of (7), and has been used in the
FIGURE 6 Actual states $x_1$, $x_2$, and their estimates, as well as unknown input $d$ and its estimates in the unbounded unknown input scenario, obtained by applying the observer in Reference 32 (Chen–Hu estimate), the observer in Reference 30 (Chak–Stan–Shre estimate) and our proposed observer (SISO estimate).

FIGURE 7 Estimation errors in the unbounded unknown input scenario for 50 different initial values (using box plots), obtained by applying the observer in Reference 32 (Chen–Hu Err.), the observer in Reference 30 (Chak–Stan–Shre Err.) and our proposed observer (SISO Err.), as well as the computed upper bounds for the state and input errors ($\delta x_k$ and $\delta d_k$).

The unknown inputs used in this example are as given in Figure 8, while the initial state estimate and noise signals (drawn uniformly) have bounds $\delta x_0^* = 0.5$, $\eta_w = 0.02$, and $\eta_v = 10^{-4}$. We also picked uniformly random coefficients, $\lambda_{ik}$, that satisfy $0 \leq \lambda_{ik} \leq 1$, $\sum_{i=1}^N \lambda_{ik} = 1$, $\forall k$. Based on the results of Theorem 2 and by solving the corresponding semi-definite programming problem, we find $P^* = \begin{bmatrix} 0.2745 & 0.1933 \\ 0.1933 & 0.4200 \end{bmatrix}$, $Y^* = \begin{bmatrix} 0.0010 \\ 0.1613 \end{bmatrix}$, and the $H_\infty$-observer gain as $L = P^*^{-1}Y^* = \begin{bmatrix} -0.3946 \\ 0.5656 \end{bmatrix}$.
7 CONCLUSION AND FUTURE WORK

We presented fixed-order set-valued $H_{\infty}$ observers for nonlinear bounded-error, discrete-time dynamic systems with unknown inputs. Necessary and sufficient Linear Matrix Inequalities for quadratic stability of the designed observer were derived for different classes of quadratically constrained nonlinear systems, including $(M, \gamma)$-QC systems, Lipschitz continuous systems, $(A, \gamma)$-QC* systems and Linear Parameter-Varying systems. Moreover, we derived additional LMI conditions and corresponding tractable semi-definite programs for obtaining the minimum $H_{\infty}$-norm for the transfer function that maps the noise signal to the state error of the stable observers.

In addition, we showed that the sequences of estimate radii of our $H_{\infty}$ observer are uniformly bounded and derived closed-form expressions for their upper bound sequences. Further, we obtained sufficient conditions for the convergence of the radii upper bound sequences and derived their steady-state values. Finally, using three illustrative examples, we demonstrated the effectiveness of our proposed design, as well as its advantages over two existing benchmark observers. For future work, we plan to generalize this framework to hybrid and switched nonlinear systems and consider other forms of CPS attacks.

CONFLICT OF INTEREST
All authors declare that they have no conflicts of interest.
DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

ENDNOTES

1 Based on the convention that the inverse of an empty matrix is an empty matrix and the assumption that operations with empty matrices are possible.

2 The readers are referred to Reference 56 (section 2) for the concise definition of uniform detectability. A spectral test can be found in Reference 57.

3 A function $\sigma : \mathbb{R}_+ \to \mathbb{R}_+$ is a $\mathcal{K}$-function if it is continuous, strictly increasing and $\sigma(0) = 0$.

4 A function $\alpha : \mathbb{R}_+ \to \mathbb{R}_+$ is a $\mathcal{K}_\infty$-function if it is a $\mathcal{K}$-function and in addition $\alpha(s) \to \infty$ as $s \to \infty$.

ORCID

Mohammad Khajenejad https://orcid.org/0000-0002-1684-8167
Sze Zhong Yong https://orcid.org/0000-0002-2104-3128

REFERENCES

1. Cárdenas A, Amin S, Sastry S. Research challenges for the security of control systems. Proceedings of the Conference on Hot Topics in Security; 2008:6:1-6:6.

2. Farwell J, Rohozinski R. Stuxnet and the future of cyber war. Survival. 2011;53(1):23-40.

3. Richards G. Hackers vs slackers. Eng Technol. 2008;3(19):40-43.

4. Slay J, Miller M. Lessons learned from the Maroochy water breach. Proceedings of the International Conference on Critical Infrastructure Protection; 2007:73-82.

5. Zetter K. Inside the cunning, unprecedented hack of Ukraine’s power grid. Wired Magazine; 2016.

6. De Nicolao G, Sparacino G, Cobelli C. Nonparametric input estimation in physiological systems: problems, methods, and case studies. Automatica. 1997;33(5):851-870.

7. Patton R, Clark R, Frank P. Fault Diagnosis in Dynamic Systems: Theory and Applications. Prentice Hall; 1989.

8. Kitanidis P. Unbiased minimum-variance linear state estimation. Automatica. 1987;23(6):775-778.

9. Gillijns S, De Moor B. Unbiased minimum-variance input and state estimation for linear discrete-time systems. Automatica. 2007;43(1):111-116.

10. Yong SZ, Zhu M, Frazzoli E. On strong detectability and simultaneous input and state estimation with a delay. Proceedings of the IEEE Conference on Decision and Control (CDC); 2015:468-475.

11. Yong SZ, Zhu M, Frazzoli E. A unified filter for simultaneous input and state estimation of linear discrete-time stochastic systems. Automatica. 2016;63:321-329.

12. Lu P, Van Kampen EJ, De Visser C, Chu Q. Framework for state and unknown input estimation of linear time-varying systems. Automatica. 2016;73:145-154.

13. Abolhasani M, Rahmani M. Robust deterministic least-squares filtering for uncertain time-varying nonlinear systems with unknown inputs. Syst Control Lett. 2018;122:1-11.

14. Yong SZ. Simultaneous input and state set-valued observers with applications to attack-resilient estimation. Proceedings IEEE American Control Conference (ACC). IEEE; 2018. p. 5167–5174.

15. Chong M, Wakaiki M, Hespanha J. Observability of linear systems under adversarial attacks. Proceedings of the American Control Conference (ACC). IEEE; 2015. p. 2439–2444.

16. Fawzi H, Tabuada P, Diggavi S. Secure estimation and control for cyber-physical systems under adversarial attacks. IEEE Trans Automat Contr. 2014;59(6):1454-1467.

17. Pasqualetti F, Dörfler F, Bullo F. Attack detection and identification in cyber-physical systems. IEEE Trans Automat Contr. 2013;58(11):2715-2729.

18. Shoukry Y, Nuzzo P, Puggelli A, et al. Imhotep-SMT: a satisfiability modulo theory solver for secure state estimation. Proceedings of the 13th International Workshop on Satisfiability Modulo Theories (SMT); 2015:3-13.

19. Kim H, Guo P, Zhu M, Liu P. Attack-resilient estimation of switched nonlinear cyber-physical systems. Proceedings of the American Control Conference (ACC); 2017;4328-4333.

20. Yong SZ, Zhu M, Frazzoli E. Resilient state estimation against switching attacks on stochastic cyber-physical systems. Proceedings of the IEEE Conference on Decision and Control (CDC); 2015:5162-5169.

21. Yong SZ, Zhu M, Frazzoli E. Switching and data injection attacks on stochastic cyber-physical systems: modeling, resilient estimation, and attack mitigation. ACM Trans Cyber-Phys Syst. 2018;2(2):9.

22. Nakahira Y, Mo Y. Dynamic state estimation in the presence of compromised sensory data. Proceedings of the IEEE Conference on Decision and Control (CDC); 2015:5808-5813; IEEE.

23. Pajic M, Tabuada P, Lee I, Pappas G. Attack-resilient state estimation in the presence of noise. Proceedings of the IEEE Conference on Decision and Control (CDC); 2015:5827-5832; IEEE.

24. Yong SZ, Foo M, Frazzoli E. Robust and resilient estimation for cyber-physical systems under adversarial attacks. Proceedings of the American Control Conference (ACC); 2016:308-315.
25. Veluvolu K, Soh Y. Discrete-time sliding-mode state and unknown input estimations for nonlinear systems. *IEEE Trans Ind Electron.* 2008;56(9):3443-3452.
26. Korbicz J, Witzczak M, Puig V. LMI-based strategies for designing observers and unknown input observers for non-linear discrete-time systems. *Bull Polish Acad Sci Techn Sci.* 2007;55(1):31-42.
27. Ha Q, Trinh H. State and input simultaneous estimation for a class of nonlinear systems. *Automatica.* 2004;40(10):1779-1785.
28. Açıklmese B, Corless M. Observers for systems with nonlinearities satisfying incremental quadratic constraints. *Automatica.* 2011;47(7):1339-1348.
29. Chakrabarty A, Corless M, Buzzard G, Žak S, Rundell A. State and unknown input observers for nonlinear systems with bounded exogenous inputs. *IEEE Trans Automat Contr.* 2017;62(11):5497-5510.
30. Chakrabarty A, Žak S, Sundaram S. State and unknown input observers for discrete-time nonlinear systems. Proceedings of the IEEE 55th Conference on Decision and Control (CDC); 2016:7111-7116; IEEE.
31. Rego B, Raffo G, Scott J, Raimondo D. Guaranteed methods based on constrained zonotopes for set-valued state estimation of nonlinear discrete-time systems. *Automatica.* 2020;111:108614.
32. Chen B, Hu G. Nonlinear state estimation under bounded noises. *Automatica.* 2018;98:159-168.
33. Chakrabarty A, Corless M. Estimating unbounded unknown inputs in nonlinear systems. *Automatica.* 2019;104:57-66.
34. Blanchini F, Sznai er M. A convex optimization approach to synthesizing bounded complexity $\ell_\infty$ filters. *IEEE Trans Automat Contr.* 2012;57(1):216-221.
35. Milanese M, Vicino A. Optimal estimation theory for dynamic systems with set membership uncertainty: an overview. *Automatica.* 1991;27(6):997-1009.
36. Khajenejad M, Yong SZ. Simultaneous input and state set-valued $H_\infty$-observers for linear parameter-varying systems. Proceedings of the American Control Conference (ACC); 2019:4521-4526; IEEE.
37. Khajenejad M, Yong SZ. Simultaneous mode, input and state set-valued observers with applications to resilient estimation against sparse attacks; 2019:1544-1550; IEEE.
38. Khajenejad M, Yong SZ. Simultaneous input and state interval observers for nonlinear systems with full-rank direct feedthrough. Proceedings of the 59th IEEE Conference on Decision and Control (CDC); 2020:5443-5448; IEEE.
39. Khajenejad M, Yong SZ. Simultaneous input and state interval observers for nonlinear systems with rank-deficient direct feedthrough. Proceedings of the European Control Conference (ECC); 2021:2311-2316; IEEE.
40. Tahir A, Açıklmese B. Synthesis of interval observers for bounded jacobian nonlinear discrete-time systems. *IEEE Control Syst Lett.* 2021;6:764-769.
41. You F, Li H, Wang F. State and unknown input simultaneous estimation for a class of nonlinear systems with time-delay. *Nonlinear Dyn.* 2016;83(3):1653-1671.
42. Khajenejad M, Jin Z, Yong SZ. Interval observers for simultaneous state and model estimation of partially known nonlinear systems; 2021:2848-2854; IEEE.
43. Wang X, Tan C, Liu L, Qi Q. A novel unknown input interval observer for systems not satisfying relative degree condition. *Int J Robust Nonlinear Control.* 2021;31(7):2762-2782.
44. Ellero N, Gucik-Derigny D, Henry D. An unknown input interval observer for LPV systems under $L_2$-gain and $L_\infty$-gain criteria. *Automatica.* 2019;103:294-301.
45. Khajenejad M, Shoaib F, Yong SZ. Interval observer synthesis for locally lipschitz nonlinear dynamical systems via mixed-monotone decompositions. Proceedings of the American Control Conference (ACC); 2022. https://arxiv.org/pdf/2202.11689.pdf.
46. Khajenejad M, Jin Z, Yong SZ. Resilient interval observer for simultaneous estimation of states, modes and attack policies. Proceedings of the American Control Conference (ACC); 2022. https://arxiv.org/pdf/2111.01873.pdf.
47. Singh K, Shen Q, Yong SZ. Mesh-based affine abstraction of nonlinear systems with tighter bounds. Proceedings of the Conference on Decision and Control (CDC); 2018:3056-3061; IEEE.
48. Gurobi Optimization, Inc. Gurobi optimizer reference manual; 2015.
49. Peters M, Iglesias P. A spectral test for observability and reachability of time-varying systems. *SIAM J Control Optim.* 1999;37(5):1330-1345.
50. Jiang Z, Wang Y. Input-to-state stability for discrete-time nonlinear systems. *Automatica.* 2001;37(6):857-869.
How to cite this article: Khajenejad M, Yong SZ. Simultaneous state and unknown input set-valued observers for quadratically constrained nonlinear dynamical systems. Int J Robust Nonlinear Control. 2022;32(12):6589-6622. doi: 10.1002/rnc.6163

APPENDIX A. PROOFS

Next, we provide proofs for our propositions, lemmas and theorems. First, for the sake of reader’s convenience, we restate a lemma from Reference 60 that we will frequently use in deriving some of our results.

Lemma 4 (60 (lemma 2.2)). Let $D$, $S$ and $F$ be real matrices of appropriate dimensions and $F^T F \preceq I$. Then, for any scalar $\varepsilon > 0$ and $x, y \in \mathbb{R}^n$, \[ 2x^T DFSy \leq \varepsilon^{-1}x^T DD^T x + \varepsilon y^T S^T Sy. \]

A.1 Proof of Proposition 2
The results follow from the facts that an inequality in $\mathbb{R}$ is preserved by multiplying the both sides by a non-negative number, or by multiplying the left hand side by a non-negative number that is not greater than 1, or by increasing the right hand side, as well as $A \preceq B \Rightarrow x^T (A - B)x \leq 0$. \hfill \Box

A.2 Proof of Proposition 3
Suppose $f_k$ is nondecreasing in its domain. Then, $2\Delta f_k^T \Delta x = \Delta f_k^T \Delta x + \Delta x^T \Delta f_k = [\Delta f_k^T \Delta x^T]^T M [\Delta f_k^T \Delta x^T]^T \succeq \gamma \geq 0$, for any $\gamma \geq 0$, with $M = \begin{bmatrix} 0_{n \times n} & I_{n \times n} \\ I_{n \times n} & 0_{n \times n} \end{bmatrix}$. Hence, $f_k$ is $(M, \gamma)$-QC. The proof for the nonincreasing case is similar. \hfill \Box

A.3 Proof of Proposition 4
Considering $M = \begin{bmatrix} -I & 0 \\ 0 & L_f \end{bmatrix}$, we have $[\Delta f_k^T \Delta q^T] M [\Delta f_k^T \Delta q^T]^T = -\Delta f_k^T \Delta f_k + L_f \Delta q^T \Delta q \succeq 0$, where the inequality is implied by the Lipschitz continuity of $f_k(\cdot)$. Hence, every Lipschitz continuous function is $\delta$-QC. To show that the converse is not true, we only need to provide a counterexample. Consider the very simple function $f(x) = x^3$ in $\mathbb{R}$. It can be easily seen that $f$ is monotonically increasing in its domain and thus, is $\delta$-QC by Proposition 3. However, $f$ is not globally Lipschitz continuous since it admits an unbounded gradient in its domain. This completes the proof. \hfill \Box

A.4 Proof of Proposition 5
By definition, $f_k(\cdot)$ is $\delta$-QC with multiplier matrix $M$ means that $[\Delta f_k^T \Delta q^T] M [\Delta f_k^T \Delta q^T]^T \succeq 0$. Then, it follows in a straightforward manner that $[\Delta f_k^T \Delta q^T] (-M)[\Delta f_k^T \Delta q^T]^T \leq \gamma$ for every $\gamma \geq 0$. \hfill \Box

A.5 Proof of Proposition 6
We observe that $[\Delta f_k^T \Delta x^T] M [\Delta f_k^T \Delta x^T]^T = -\Delta f_k^T \Delta f_k \geq -L_f^2 \|\Delta x\|^2 \geq -L_f^2 (2r)^2 = -4r^2 L_f^2$, where the second and third inequalities hold by Lipschitz continuity of $f(\cdot)$ and boundedness of the state space, respectively. \hfill \Box

A.6 Proof of Proposition 7
It follows from Definitions 5 and 6 and $\gamma \geq 0$ that \[ -\Delta f_k^T \Delta f_k + 2\Delta f_k^T A \Delta x - \Delta x^T A^T A \Delta x \succeq \gamma \geq 0. \]
This, along with Lemma 4, imply that

\[(1 - \epsilon)\|\Delta f_k\|^2 \leq (\epsilon^{-1} - 1)\Delta x^T A^T A \Delta x, \quad \forall \epsilon > 0\]

\[\Rightarrow \|\Delta f_k\|^2 \leq \frac{1}{\epsilon} \Delta x^T A^T A \Delta x, \quad \forall 0 < \epsilon < 1.\]

By taking the limit of the both sides when \(\epsilon \to 1\) and by the continuity of the “\(\leq\)” operator, we obtain

\[\|\Delta f_k\|^2 \leq \Delta x^T A^T A \Delta x \leq \lambda_{\text{max}}(A^T A)\|\Delta x\|^2.\]

Finally, the result is obtained by taking the square root of both sides of the above inequality.

\[\square\]

A.7 Proof of Lemma 1

First, notice that \(\Delta f_k = A \Delta x + \Delta g\). Given this and \(\|g(x)\| \leq r\), we can conclude that

\[
\begin{bmatrix}
\Delta f_k^T \\
\Delta x^T
\end{bmatrix}
M
\begin{bmatrix}
\Delta f_k \\
\Delta x
\end{bmatrix}^T
= -\Delta f_k^T \Delta f_k + 2 \Delta x^T A \Delta f_k - \Delta x^T A \Delta x
= -(\Delta f_k - A \Delta x)^T (\Delta f_k - A \Delta x) = -\Delta g^T \Delta g \geq -(2r)^2.
\]

\[\square\]

A.8 Proof of Proposition 8

By construction, we have the following condition:

\[
M - \begin{bmatrix}
-I_{N,M} & A \\
A^T & -A^T A
\end{bmatrix}
= \begin{bmatrix}
M_{11} + I & 0 \\
0 & M_{22} + M_{12} M_{12}
\end{bmatrix}
\preceq 0,
\]

since both submatrices on the diagonal are negative semi-definite by assumption.

\[\square\]

A.9 Proof of Proposition 9

The global Lipschitz continuity of LPV systems can be shown as follows:

\[
\Delta f_k \triangleq \|f_k(x_1) - f_k(x_2)\| = \left\| \sum_{i=1}^{N} \lambda_{i,k} A^i \Delta x_k \right\| \\
\leq \sum_{i=1}^{N} \lambda_{i,k} \|A^i\| \|\Delta x_k\| \\
\leq \|A^m\| \|\Delta x_k\|,
\]

with \(\|A^m\| = \max_{i=1 \ldots N} \|A^i\|\), where the first and second inequalities hold by submultiplicative inequality for norms and positivity of \(\lambda_{i,k}\), while the third inequality holds by the facts that \(0 \leq \lambda_{i,k} \leq 1\) and \(\sum_{i=1}^{N} \lambda_{i,k} = 1\).

\[\square\]

A.10 Proof of Lemma 2

Aiming to derive the governing equation for the evolution of the state errors, from (12) and (15), we obtain

\[
\hat{d}_{1,k} = M_1(C_1 \tilde{x}_{1,k} + \Sigma d_{1,k} + v_{1,k}).
\]  

(A1)

Moreover, from (3), (12) and (16)–(19), we have

\[
\hat{d}_{2,k-1} = M_2[C_2(\Delta f_{k-1} + G_1 \tilde{d}_{1,k-1} + G_2 d_{2,k-1} + W w_{k-1}) + v_{2,k}],
\]  

(A2)

and by plugging \(M_1 = \Sigma^{-1}\) into (A1), we obtain

\[
\tilde{d}_{1,k} = \hat{d}_{1,k} - \hat{d}_{1,k} = -M_1(C_1 \tilde{x}_{1,k} + v_{1,k}).
\]  

(A3)
where $\Delta f_k \triangleq f_k(x_k) - f_k(\hat{x}_k)$. Then, by setting $M_2 = (C_2G_2)^\dagger$ in (A2) and using (A3), we have

$$\ddot{d}_{2,k-1} = -M_2[C_2(\Delta f_{k-1} - G_1M_2(C_1\ddot{x}_{k-1|k-1} + v_{1,k-1}) + Ww_{k-1}) + v_{2,k}].$$

(A4)

Furthermore, it follows from (3), (18), and (19) that

$$\ddot{x}_{k|k}^* = \Delta f_{k-1} + G_1\ddot{d}_{1,k-1} + G_2\ddot{d}_{2,k-1} + Ww_{k-1}.$$  

(A5)

In addition, by plugging $\ddot{d}_{k-1}$ and $\ddot{d}_{k-2}$ from (A3) and (A4) into (A5), by (12) and (20), we obtain

$$\ddot{x}_{k|k}^* = (I - \tilde{L}C_2)^\dagger \dot{\bar{v}}_k.$$  

(A6)

$$\ddot{x}_{k|k}^* = \Phi[\Delta f_{k-1} - G_1M_1C_1\dot{x}_{k-1|k-1}] + \bar{w}_k.$$  

(A7)

where $\bar{v}_k \triangleq v_{2,k}$, $\bar{w}_k \triangleq -\Phi(G_1M_1v_{1,k-1} - Ww_{k-1}) - G_2M_2v_{2,k}$, and $\Phi \triangleq I - G_2M_2C_2$. Finally, combining (A6) and (A7) returns the results.

**A.11 Proof of Theorem 1**

First, note that the state error dynamics (21) without bounded noise signals $w_k$ and $v_k$ can be rewritten as

$$\ddot{x}_{k+1|k+1} = \ddot{A}\ddot{x}_{k|k} + \ddot{B}\Delta f_k,$$

with $\ddot{A} \triangleq -(I - LC_2)\Phi\Psi$ and $\ddot{B} \triangleq (I - LC_2)\Phi$. Moreover, considering a quadratic positive definite candidate Lyapunov function

$$V_k^{\text{run}} = \dddot{x}_{k|k}^T P \dddot{x}_{k|k},$$

(A9)

with $P > 0$, (22) is equivalent to:

$$(\ddot{A}\ddot{x}_{k|k} + \ddot{B}\Delta f_k)^T P(\ddot{A}\ddot{x}_{k|k} + \ddot{B}\Delta f_k) \leq (1 - \alpha)\dddot{x}_{k|k}^T P \dddot{x}_{k|k},$$

which can be reorganized as:

$$\begin{bmatrix} \Delta f_k \\ \ddot{x}_{k|k} \end{bmatrix}^T \begin{bmatrix} B^T \dddot{B} & B^T \dddot{P} \dddot{A} \\ \dddot{A}^T \dddot{B} & \dddot{A}^T \dddot{P} \dddot{A} + (\alpha - 1)P \end{bmatrix} \begin{bmatrix} \Delta f_k \\ \ddot{x}_{k|k} \end{bmatrix} \leq 0.$$  

(A10)

Then, for Class 0 functions (i.e., $(M, \gamma)$-QC functions), by Definition 5, we have

$$\begin{bmatrix} \Delta f_k \\ \ddot{x}_{k|k} \end{bmatrix}^T \begin{bmatrix} -M_{11} & -M_{12} \\ -M_{21}^T & -M_{22} \end{bmatrix} \begin{bmatrix} \Delta f_k \\ \ddot{x}_{k|k} \end{bmatrix} \leq -\gamma.$$  

(A11)

In other words, we want (A10) to hold for any pair of $(\ddot{x}_{k|k}, \Delta f_k)$ that satisfy (A11). By applying $S$-procedure, this is equivalent to the existence of $\kappa \geq 0$, such that:

$$\kappa \begin{bmatrix} -M_{11} & -M_{12} & 0 \\ -M_{21}^T & -M_{22} & 0 \\ 0 & 0 & \gamma \end{bmatrix} \begin{bmatrix} B^T \dddot{B} & B^T \dddot{P} \dddot{A} \\ \dddot{A}^T \dddot{B} & \dddot{A}^T \dddot{P} \dddot{A} + (\alpha - 1)P \end{bmatrix} \begin{bmatrix} -M_{11} & -M_{12} \\ -M_{21}^T & -M_{22} \\ 0 & 0 & \gamma \end{bmatrix} \geq 0.$$  

(A12)

Since $\gamma \geq 0$ by assumption, the above is equivalent to:

$$\exists \kappa \geq 0 : \kappa \begin{bmatrix} -M_{11} & -M_{12} \\ -M_{21}^T & -M_{22} \end{bmatrix} - \begin{bmatrix} B^T \dddot{B} & B^T \dddot{P} \dddot{A} \\ \dddot{A}^T \dddot{B} & \dddot{A}^T \dddot{P} \dddot{A} + (\alpha - 1)P \end{bmatrix} \geq 0.$$
Now, note that \( \kappa = 0 \) is not a valid choice, since if \( \kappa = 0 \), it can be shown that there is no \( P > 0 \) that satisfies the above inequality. Hence, \( \kappa > 0 \), and we equivalently obtain \( \exists P > 0, \kappa > 0 \), such that

\[
\begin{bmatrix}
\hat{B}^T PB & \hat{B}^T PA \\
\hat{A}^T PB & \hat{A}^T PA + (a-1)P
\end{bmatrix} \leq \kappa \begin{bmatrix}
-M_{11} & -M_{12} \\
-M_{12} & -M_{22}
\end{bmatrix} \iff \begin{bmatrix}
-\kappa M_{11} - \hat{B}^T PB & -\kappa M_{12} - \hat{B}^T PA \\
-\kappa M_{12} - \hat{A}^T PB & -\kappa M_{22} - \hat{A}^T PA + (1-a)P
\end{bmatrix} \geq 0,
\]

which is, in turn, equivalent to:

\[
\forall \epsilon > 0, \exists P > 0, \kappa > 0 : 
\begin{bmatrix}
-\kappa M_{11} - \hat{B}^T PB + \epsilon I & -\kappa M_{12} - \hat{B}^T PA \\
-\kappa M_{12} - \hat{A}^T PB & -\kappa M_{22} - \hat{A}^T PA + (1-a)P + \epsilon I
\end{bmatrix} > 0.
\]

Then, by applying Schur complement, we obtain

\[
\forall \epsilon > 0 : \begin{cases}
\hat{M}_{22} > \hat{M}_{12}, \\
-\kappa M_{11} - \hat{B}^T PB + \epsilon I > 0,
\end{cases}
\]

(A13)

with \( \hat{M}_{22} \triangleq -\kappa M_{22} - \hat{A}^T PA + (1-a)P + \epsilon I \) and \( \hat{M}_{12} \triangleq (\kappa M_{12}^T + \hat{A}^T PB)(-\kappa M_{11} - \hat{B}^T PB + \epsilon I)^{-1}(\kappa M_{12} + \hat{B}^T PA) \), where the second inequality is equivalent to:

\[
-\kappa M_{11} - \hat{B}^T PP^{-1} PB \geq 0 \iff \begin{bmatrix}
P \\
\hat{B}^T P
\end{bmatrix} \geq 0.
\]

On the other hand, \( \forall \epsilon > 0 : \hat{M}_{22} > \hat{M}_{12} \) is equivalent to:

\[
\forall \epsilon > 0, \exists \Gamma \geq 0 : 
\begin{cases}
\Gamma \geq \hat{M}_{12} \geq 0, \\
-\kappa M_{22} - \hat{A}^T PP^{-1} PA + (1-a)P + \epsilon I > \Gamma.
\end{cases}
\]

(A14)

Since the second inequality holds for all \( \epsilon \), it is equivalent to \( -\kappa M_{22} - \hat{A}^T PP^{-1} PA + (1-a)P \geq \Gamma \), which, by applying Schur complement, is equivalent to:

\[
\begin{bmatrix}
P \\
\hat{A}^T P
\end{bmatrix} \geq 0.
\]

(A15)

By also applying Schur complement to (A14), we obtain

\[
\forall \epsilon > 0 : \Gamma \geq \hat{M}_{12},
\]

\[
\iff \forall \epsilon > 0 : \begin{bmatrix}
-\kappa M_{11} - \hat{B}^T PB + \epsilon I & \kappa M_{12} + \hat{B}^T PA \\
\kappa M_{12} + \hat{A}^T PB & \Gamma
\end{bmatrix} \geq 0,
\]

\[
\iff \forall \epsilon > 0 : \begin{bmatrix}
\Gamma + \delta I & \kappa M_{12} + \hat{A}^T PB \\
\kappa M_{12} + \hat{B}^T PA + (e + \delta) I
\end{bmatrix} \geq 0.
\]

Applying Schur complement to the above yields
\[ \forall \epsilon, \delta > 0 : \\
\tilde{\Gamma} + \delta I > 0 \iff \tilde{\Gamma} \geq 0 \land -\kappa M_{11} - \tilde{B}^T P \tilde{B} + (\epsilon + \delta)I > \Xi, \]

where \( \Xi \triangleq (\kappa M_{12} + \tilde{B}^T P \tilde{A})(\tilde{\Gamma} + \delta I)^{-1}(\kappa M_{12}^T + \tilde{A}^T P \tilde{B}) \). Then, from the last inequality, we equivalently obtain: \( \forall \epsilon, \delta > 0 : \exists \tilde{Q} \)
such that:

\[ -\kappa M_{11} - \tilde{B}^T P \tilde{B} + (\epsilon + \delta)I > \tilde{Q} \iff -\kappa M_{11} - \tilde{B}^T P \tilde{B} \geq \tilde{Q} \iff \begin{bmatrix} P & \tilde{P} \\ \tilde{B}^T P & -\kappa M_{11} - \tilde{Q} \end{bmatrix} \geq 0 \quad (A16) \]

and \( \tilde{Q} \geq (\kappa M_{12} + \tilde{B}^T P \tilde{A})(\tilde{\Gamma} + \delta I)^{-1}(\kappa M_{12}^T + \tilde{A}^T P \tilde{B}) \), which is equivalent to

\[ \forall \delta > 0 : \begin{bmatrix} \tilde{\Gamma} + \delta I & -\kappa M_{12}^T - \tilde{A}^T P \tilde{B} \\ -\kappa M_{12} - \tilde{B}^T P \tilde{A} & \tilde{Q} \end{bmatrix} \geq 0. \quad (A17) \]

Furthermore, pre- and post-multiplication of \((A17)\) by \( \begin{bmatrix} I & 0 \\ 0 & \Psi^T \end{bmatrix} \) and \( \begin{bmatrix} I & 0 \\ 0 & \Psi \end{bmatrix} \), as well as the fact that \(-B\Psi = \tilde{A}\) by definition and applying Schur complement, lead to

\[ \forall \delta > 0 : \begin{bmatrix} \tilde{\Gamma} + \delta I & -\kappa M_{12}^T \Psi + \tilde{A}^T P \tilde{A} \\ -\Psi^T \kappa M_{12} + \tilde{A}^T P \tilde{A} & \Psi^T \tilde{Q} \Psi \end{bmatrix} \geq 0 \iff \forall \delta > 0 : \Psi^T \tilde{Q} \Psi \geq \tilde{Q}, \]

where \( \tilde{Q} \triangleq (-\Psi^T \kappa M_{12} + \tilde{A}^T P \tilde{A})(\tilde{\Gamma} + \delta I)^{-1}(-\kappa M_{12}^T \Psi + \tilde{A}^T P \tilde{A}) \). Equivalently, we can represent this as:

\[ \forall \delta > 0, \exists \tilde{Z} : \tilde{Z} \geq -\kappa M_{12}^T \Psi + \tilde{A}^T P \tilde{A}, \Psi^T \tilde{Q} \Psi \geq \tilde{Z}^T(\tilde{\Gamma} + \delta I)^{-1}\tilde{Z}, \]

where the first inequality is equivalent to:

\[ \begin{bmatrix} P & \tilde{P} \\ \tilde{A}^T P & \tilde{Z} + \kappa M_{12}^T \Psi \end{bmatrix} \geq 0, \]

and the second inequality can be written as:

\[ \forall \delta > 0 : \begin{bmatrix} \tilde{\Gamma} + \delta I & \tilde{Z} \\ \tilde{Z}^T & \Psi^T \tilde{Q} \Psi \end{bmatrix} \geq 0 \iff \begin{bmatrix} \tilde{\Gamma} & \tilde{Z} \\ \tilde{Z}^T & \Psi^T \tilde{Q} \Psi \end{bmatrix} \geq 0, \quad (A18) \]

where the forward direction (i.e., \( \Rightarrow \)) follows from the fact that the limit of a convergent sequence of positive semi-definite matrices is a positive semi-definite matrix, while the backward direction (i.e., \( \Leftarrow \)) holds since the summation of two positive semi-definite matrices is positive semi-definite. Finally, by defining \( Y \triangleq \tilde{P} \tilde{L} \) with \( \tilde{\Gamma} > 0 \) and the LMIs in \((A13)-(A16)\) and \((A18)\), we obtain the results for Class 0 functions.

Next, we consider system classes I,II–III by appealing to the fact that they are Class 0 systems with additional information about \( M_{11}, M_{12} \) and \( M_{22} \):

- Case I: From Proposition 4, we can replace \( M_{11}, M_{12}, \) and \( M_{22} \) with \(-I, 0 \) and \( L^2 I \).
- Case II: By Definition 6, we can substitute \( M_{11}, M_{12}, \) and \( M_{22} \) with \(-I, A \) and \(-A^T A\).
- Case III: Note that by Proposition 9, an LPV function is Lipschitz continuous by Lipschitz constant \( \hat{\sigma}_m \triangleq \max_{e \in E} \| A^e \| \). Hence, we can apply Case I by replacing \( L_f \) with \( \hat{\sigma}_m \).

Finally, suppose \( \gamma < 0 \). Then, the fact that \( \kappa = 0 \) is not a valid choice (as previously discussed) implies that \( \kappa \gamma < 0 \) and hence, the matrix in the left hand side of \((A12)\) always contains a strictly negative diagonal element \( \kappa \gamma < 0 \) and consequently, \((A12)\) can never hold, and thus, neither can \((A10)\).
A.12 Proof of Proposition 10
To prove the necessity of (28), we use contraposition. Suppose that the LMIs in (28) are feasible. Then, we will show that there exists a candidate Lyapunov function $\tilde{V}_{wn}^k = \tilde{x}_{k|k}^T \tilde{P}_{\lambda k}$, for some $\tilde{P} > 0$, such that $\Delta \tilde{V}_{wn}^k \triangleq \tilde{V}_{wn}^{k+1} - \tilde{V}_{wn}^k > 0$ and hence, by the Lyapunov instability theorem (62 (theorem 3.3)), the error system is unstable. Therefore, the conditions in (28) are necessary for the stability of the observer. To do so, first note that

$$\Delta \tilde{V}_{wn}^k = \Delta f_k^T \Phi^T (I - \tilde{L}C_2)^T \tilde{P} (I - \tilde{L}C_2) \Phi \Delta f_k + \tilde{x}_{k|k}^T \Psi^T (I - \tilde{L}C_2)^T \tilde{P} (I - \tilde{L}C_2) \Phi \Psi - \tilde{P} \tilde{x}_{k|k}$$

$$- 2 \Delta f_k^T \Phi^T (I - \tilde{L}C_2)^T \tilde{P} (I - \tilde{L}C_2) \Phi \Psi \tilde{x}_{k|k}$$

(A19)

Then, (28), along with setting $\tilde{L} = \tilde{P}^{-1} \tilde{Y} = \tilde{Y} = \tilde{P} \tilde{L}$, defining $\tilde{S} \triangleq \tilde{P} - C_2^T \tilde{Y}^T - \tilde{Y} C_2$, $\tilde{\Pi} \succ 0$ and applying Schur complement, result in

$$0 < \tilde{L}^T \tilde{P} \tilde{L} \preceq \tilde{I}$$

(A20)

as well as

$$\Delta \tilde{V}_{wn}^k = \Delta f_k^T \Phi^T (\tilde{S} + C_2^T \tilde{P} L C_2) \Phi \Delta f_k + \tilde{x}_{k|k}^T \Psi^T \Phi^T (\tilde{S} + C_2^T \tilde{P} L C_2) \Phi \Psi - \tilde{P} \tilde{x}_{k|k}$$

$$- 2 \Delta f_k^T \Phi^T \tilde{S} \Phi \Psi \tilde{x}_{k|k} - 2 \Delta f_k^T \Phi^T C_2^T \tilde{P} L C_2 \Phi \Psi \tilde{x}_{k|k}$$

(A21)

Then, (A20), (A21), and Lemma 4 imply that there exists $\tilde{\eta} > 0$ such that

$$\Delta \tilde{V}_{wn}^k \geq \Delta f_k^T \Phi^T (\tilde{S} + \tilde{\eta} C_2^T C_2) \Phi \Delta f_k + \tilde{x}_{k|k}^T \Psi^T \Phi^T (\tilde{S} + \tilde{\eta} C_2^T C_2) \Phi \Psi - \tilde{P} \tilde{x}_{k|k}$$

$$- 2 \Delta f_k^T \Phi^T \tilde{S} \Phi \Psi \tilde{x}_{k|k} - 2 \Delta f_k^T \Phi^T C_2^T \tilde{S} \Phi \Psi C_2 \Phi \Psi \tilde{x}_{k|k} \triangleq \Delta \tilde{V}_{wn}^k$$

(A22)

As with the proof of Theorem 1, we first consider Class 0 systems, where by plugging $\tilde{\Pi}_{11}$, $\tilde{\Pi}_{12}$, and $\tilde{\Pi}_{22}$ given in (29) into $\tilde{\Pi}$, as defined in (28), we obtain:

$$0 < \left[ \Delta f_k^T \tilde{x}_{k|k}^T \right] \tilde{\Pi} \left[ \Delta f_k^T \tilde{x}_{k|k}^T \right]^T = \Delta \tilde{V}_{wn}^k - \delta_f$$

(A23)

where

$$\delta_f \triangleq \Delta f_k^T M_{11} \Delta f_k + \Delta f_k^T M_{12} \Delta f_k + \tilde{x}_{k|k}^T M_{22} \tilde{x}_{k|k} = \left[ \Delta f_k^T \tilde{x}_{k|k}^T \right] \mathcal{M} \left[ \Delta f_k^T \tilde{x}_{k|k}^T \right]^T \geq \gamma \geq 0.$$  

(A24)

Finally, from (A22), we have $\Delta \tilde{V}_{wn}^k \geq \Delta \tilde{V}_{wn}^k > \delta_f \geq 0 \Rightarrow \Delta \tilde{V}_{wn}^k > 0$.

Furthermore, the proof for system classes I, II–II can be obtained from the above result for Class 0 with suitable values of $M_{11}$, $M_{12}$ and $M_{22}$ as discussed in the proof of Theorem 1.

A.13 Proof of Lemma 3
To show that uniform detectability is sufficient for existence of an observer, notice that for a Class III function $f_k(\cdot)$, (21) can be written as

$$\tilde{x}_{k|k} = (I - \tilde{L}C_2)\tilde{A}_{k-1} \tilde{x}_{k-1|k-1} + (I - \tilde{L}C_2)\tilde{W}_{k-1} - \tilde{L}\tilde{V}_{k-1}.$$  

(A25)

where

$$\tilde{W}_{k-1} \triangleq -(I - G_2 M_2 C_2)(G_1 M_1 v_{1|k-1} - w_{k-1}) - G_2 M_2 v_{2,k},$$

$$\tilde{A}_{k} \triangleq \Phi \left( \sum_{i=1}^{N} \lambda_{i,k} A^i - \Psi \right),$$

$$\tilde{V}_{k-1} \triangleq v_{2,k}.$$
Now, consider the following linear time-varying system without unknown inputs:

\[
x_{k+1} = \bar{A}_k x_k + \bar{w}_k, y_k = C_2 x_k + \bar{v}_k.
\] (A26)

Systems (A25) and (A26) are equivalent from the viewpoint of estimation, since the estimation error equations for both problems are the same, hence they both have the same objective. Therefore, the pair \((\bar{A}_k, C_2)\) needs to be uniformly detectable such that the observer is stable (section 5).

Moreover, as for the necessity of the strong detectability of the constituent LTI systems, suppose for contradiction, that there exists a stable observer for system (3) with any sequence \(\{\lambda_{i,k}\}_{k=0}^{\infty}\) for all \(i \in \{1, 2, \ldots, N\}\) that satisfies \(0 \leq \lambda_{i,k} \leq 1, \sum_{i=1}^{N} \lambda_{i,k} = 1, \forall k\), but one of the constituent linear time-invariant systems (e.g., \((A', G, C, H)\)) is not strongly detectable. Since the observer exists for any sequence of \(\lambda_{i,k}\), that means that an observer also exists when \(\lambda_{i,k} = 1\) and \(\lambda_{i,k} = 0, \forall i \neq j\) for all \(k\). However, we know from Reference 14 that strong detectability is necessary for the stability of the linear time-invariant system \((A', G, C, H)\), which is a contradiction. Hence, the proof is complete.

A.14 Proof of Theorem 2

We use a similar approach as in the proof of Theorem 1. First, consider the error dynamics with bounded noise signals (21) and the candidate Lyapunov function \(V^n_k \triangleq \bar{x}^T_{k|k} P \bar{x}_{k|k}\). Observe that

\[
\Delta V^n_k \triangleq V^n_{k+1} - V^n_k = \Delta V^{wn}_k + \Delta r_k,
\] (A27)

where \(V^{wn}_k\) is the Lyapunov function for the error dynamics without noise signals, defined in (A9), and

\[
\Delta r_k \triangleq 2(\Delta f^T_k - \bar{x}^T_{k|k} \Psi^T) (I - \tilde{L} C_2) P \bar{W}(\tilde{L}) \bar{w}_k + \bar{w}^T_k \bar{W}(\tilde{L})^T P \bar{W}(\tilde{L}) \bar{w}_k,
\] (A28)

with \(\Phi, \Psi, \bar{w}_k\) and \(\bar{W}(\tilde{L})\) defined in Lemma 2. We will show that for each system class,

\[
\Delta r_k \triangleq \Delta r_k - \rho^2 \bar{w}^T_k \bar{w}_k + \bar{x}^T_k (I - \alpha P) \bar{x}_k \leq 0.
\] (A29)

Then, by (A27) and (A29) in addition to the fact that \(\Delta V^{wn}_k \leq -\alpha \bar{x}^T_k P \bar{x}_k\) (follows from Theorem 1), we have

\[
\Delta V^n_k \leq \rho^2 \bar{w}^T_k \bar{w}_k - \bar{x}^T_{k|k} \bar{x}_{k|k}.
\] (A30)

Summing up both sides of (A30) from zero to infinity, returns \(V^{wn}_0 - V^n_0 \leq \rho^2 \sum_{k=0}^{\infty} \bar{w}^T_k \bar{w}_k - \sum_{k=0}^{\infty} \bar{x}^T_k \bar{x}_{k|k} - \rho^2 \sum_{k=0}^{\infty} \bar{w}^T_k \bar{w}_k - \sum_{k=0}^{\infty} \bar{x}^T_k \bar{x}_{k|k}\), where at each time step \(k\), \(\bar{w}_k = [w^T_k \quad v^T_k]^T\). Then, it follows from setting the initial conditions to zero that

\[
\sum_{k=0}^{\infty} \bar{x}^T_k \bar{x}_{k|k} \leq \rho^2 \sum_{k=0}^{\infty} \bar{w}^T_k \bar{w}_k.
\]

Thus, it remains to show that (A29) holds for each system class 0, I, II–III. Plugging the expression for \(\bar{W}(\tilde{L})\) from Lemma 2 into (A28), we obtain

\[
\Delta r_k = \bar{x}^T_{k|k} (I - \alpha P) \bar{x}_{k|k} + 2(\Delta f^T_k - \bar{x}^T_{k|k} \Psi^T) (I - \tilde{L} C_2) P \bar{W}(\tilde{L}) \bar{w}_k
\]

\[
+ \bar{w}^T_k (R^T PR - 2\Omega^T \bar{W}(\tilde{L})^T P \bar{W}(\tilde{L}) \bar{w}_k - \Omega^T \bar{W}(\tilde{L})^T P \bar{W}(\tilde{L}) \bar{w}_k - \rho^2 I) \bar{w}_k,
\] (A31)

On the other hand, note that \(\Pi \succ 0 \Leftrightarrow I - \Gamma \succ 0\) and \(\begin{bmatrix} \Gamma & Y^T \\ Y & P \end{bmatrix} \succ 0\), which by pre and postmultiplication by \(\begin{bmatrix} I & 0 \\ 0 & P^{-1} \end{bmatrix}\) and the fact that \(Y = PL\), is equivalent to \(I - \Gamma \succ 0\) and \(\begin{bmatrix} \Gamma & L \\ L & P^{-1} \end{bmatrix} \succ 0\). Applying Schur complement to the latter, \(\Pi \succ 0\) is equivalent to

\[
0 \leq L^T PL \leq \Gamma \leq I.
\] (A32)

Now, (A31), (A32) and Lemma 4, imply that:

\[
\Delta r_k \leq \bar{w}^T_k (R^T PR - 2R^T Y \Omega + \Omega^T \Gamma \Omega + (\epsilon_1^2 + \epsilon_2^2) \Omega^T \Omega - \rho^2 I) \bar{w}_k + \bar{x}^T_{k|k} (I - \alpha P + \epsilon_1 \Psi^T \Phi^T C_2 C_2 \Phi \Psi) \bar{x}_{k|k}
\]

\[
+ 2\Delta f^T_k \Phi^T (PR - Y \Omega - C_2^T Y^T R) \bar{w}_k - 2\bar{x}^T_{k|k} \Psi^T \Phi^T (PR - Y \Omega - C_2^T Y^T R) \bar{w}_k + \epsilon_2 \Delta f^T_k (\Phi^T C_2 C_2 \Delta f^T_k) \Delta f_k \triangleq \bar{\Delta} r_k.
\] (A33)
We first consider Class 0 systems, where the fact that \( f_k(\cdot) \) is \((\mathcal{M}, \gamma)\)-QC with \( \gamma \geq 0 \) implies that
\[
-\Delta_k^T M_{11} \Delta_k - x_{k|k}^T M_{22} \tilde{x}_{k|k} - 2 \Delta_k^T M_{12} \tilde{x}_{k|k} \triangleq \tilde{S} \leq -\gamma \leq 0,
\]
which, in addition to (A33), return \( \Delta_k \leq \tilde{\Delta}_k + \tilde{S} \leq 0 \), where \( \tilde{\Delta}_k \triangleq \begin{bmatrix} \tilde{w}_k^T & \tilde{x}_k^T & \Delta_k^T \end{bmatrix}^T \) and \( \mathcal{N} \) is the matrix in (32) with its elements defined in (33) and (34). Finally, we can obtain the results for system classes I, II–III with suitable values of \( M_{11}, M_{12} \) and \( M_{22} \) as described in the proof of Theorem 1.

### A.15 Proof of Theorem 3
Consider the noisy state error dynamics system in (21), where we treat the augmented noise signal \( \tilde{w}_k \) as an external input to the system. Recall from the proof of Theorem 2 that as a result of \( H_\infty \) observer design, the Lyapunov function \( V_k \triangleq \tilde{x}_{k|k}^T P \tilde{x}_{k|k} \) satisfies (A30), which is equivalent to \( \Delta V_k \leq -\alpha_3(||\tilde{x}_{k|k}||) + \sigma(||\tilde{w}_k||) \), with the \( K_\infty \)-function \( \alpha_3(r) \triangleq r^2 \) and the \( K \)-function \( \sigma(s) \triangleq \eta s^2 \). This implies that \( V_k \) is an ISS-Lyapunov function for system (21) (cf. Reference 58 (definition 3.2)), that is, (21) admits an input-to-state stable (ISS)-Lyapunov function. Equivalently, (21) is ISS and UBIBS, and admits a \( K \)-asymptotic gain by Reference 58 (theorem 1).

### A.16 Proof of Theorem 4
In this theorem, we aim to find two upper bounds of \( ||\tilde{x}_{k|k}|| \) such that \( ||\tilde{x}_{k|k}|| \leq \tilde{\delta}_{k,1} \) and \( ||\tilde{x}_{k|k}|| \leq \tilde{\delta}_{k,2} \). To derive the first upper bound, notice that it follows from (A30) that
\[
\tilde{x}_{k|k}^T P \tilde{x}_{k|k} \leq \rho^2 \tilde{w}_k^T \tilde{w}_k + \tilde{x}_{k-1|k-1}^T (P - I) \tilde{x}_{k-1|k-1},
\]
and by applying Rayleigh’s inequality, we obtain
\[
\lambda_{\min}(P)||\tilde{x}_{k|k}||^2 \leq \rho^2 \tilde{w}_k^T \tilde{w}_k + \lambda_{\max}(P - I)||\tilde{x}_{k-1|k-1}||^2
\]
\[
\Rightarrow ||\tilde{x}_{k|k}|| \leq \sqrt{\frac{\rho^2 \tilde{w}_k^T \tilde{w}_k + \lambda_{\max}(P - I)||\tilde{x}_{k-1|k-1}||^2}{\lambda_{\min}(P)}}.
\]
Repeating this procedure \( k \) times and considering the fact that \( \lambda_{\max}(P - I) = \lambda_{\max}(P) - 1 \) as a consequence of Weyl’s Theorem, we obtain \( ||\tilde{x}_{k|k}|| \leq \tilde{\delta}_{k,1} \) with \( \tilde{\delta}_{k,1} \) given in (38) and \( \theta_1 = \frac{\lambda_{\max}(P - I)}{\lambda_{\max}(P)} \).

Next, we find the second upper bound for \( ||\tilde{x}_{k|k}|| \leq \tilde{\delta}_{k,2} \) as a consequence of Weyl’s Theorem (4.3.1), we obtain \( \tilde{x}_{k|k} \leq \tilde{x}_{k,1} \) with \( \tilde{x}_{k,1} \) given in (38) and \( \theta_1 = \frac{\lambda_{\max}(P - I)}{\lambda_{\max}(P)} \).

(I) If \( f_k(\cdot) \) is a Class I function, then, the result in (36) with \( \theta \) defined in (41), directly follows from Lipschitz continuity of \( f_k(\cdot) \), as well as applying triangle and submultiplicative inequalities for norms on (21). Moreover, the result in (37) with \( \beta \) defined in (41), is obtained by triangle and submultiplicative inequalities, (10), (A3), and (A4).

(II) If \( f_k(\cdot) \) is a Class II function, then by Proposition 7, it is a Class I function with \( L_f = \sqrt{\lambda_{\max}(\mathcal{A}A^T \mathcal{A})} \). The rest of the proof is similar to the proof for Class I.

(III) If \( f_k(\cdot) \) is a Class III function, we first find closed-form expressions for the state and input estimation errors through the following lemma.

**Lemma 5.** The state and input estimation errors are
\[
\tilde{x}_{k|k} = \sum_{i=1}^{k-2} \sum_{j=0}^{k-2} A_{e,k-j}(\tilde{x}_{k-1|k-1} - \tilde{L}\tilde{y}_{k-1}) + \sum_{j=0}^{k-1} A_{e,k-j}(\tilde{x}_{k|k}, 0),
\]
\[
\tilde{d}_{k-1} = -\sum_{i=1}^{\mathcal{N}} \lambda_i \tilde{x}_{k|k} - (V_1 M_1 C_1 + V_2 M_2 C_2 A_{e,c}) \tilde{x}_{k|k} - (V_2 M_2 C_2 G_1 M_1 - V_1 M_1) T_1 \tilde{y}_{k-1} - V_2 M_2 C_2 \tilde{w}_{k-1} - V_2 M_2 T_2 \tilde{v}_k.
\]

**Proof.** Starting from (A25) and applying simple induction return the results for the state errors. Then, the expression for the input errors follows from (A3), (A4) and (10).

Now, we are ready to show that \( ||\tilde{x}_{k|k}|| \leq \tilde{\delta}_{k,2} \triangleq \delta_{0,2}^\infty + \tilde{\eta} \sum_{i=1}^{k-1} \theta_2^{-i} \) for LPV (Class III) functions. First, we define
\( B_{e,k} \triangleq \prod_{j=0}^{k-1} A_{e,k-j} \),
\( C_{e,k}^i \triangleq \prod_{j=0}^{i-2} A_{e,k-j} \),
\( \tilde{t}_k \triangleq \Psi \tilde{w}_k - \tilde{v}_k \),
(A34)

for \( 1 \leq i \leq k \). Then, from Lemma 5, we have
\[
\| \tilde{x}_{k|k} \| \leq \| B_{e,k} \| \| \tilde{x}_{0|0} \| + \sum_{i=1}^{k} C_{e,k}^i \| \tilde{t}_{k-i} \| ,
\]
(A35)

by triangle inequality and submultiplicativity of norms. Moreover, by a similar reasoning, we find
\[
\| B_{e,k} \| \leq \prod_{j=0}^{k-1} \sum_{i=1}^{N} \lambda^i_{k-j} \Psi \Phi (A^i - G_1 M_1 C_1) \| \leq \theta_2^k ,
\]
\[
\| \sum_{i=1}^{k} C_{e,k}^i \tilde{t}_{k-i} \| \leq \sum_{i=1}^{k} \| C_{e,k}^i \| \| \tilde{t}_{k-i} \| ,
\]
\[
\| C_{e,k}^i \| \leq \prod_{j=0}^{i-2} \sum_{s=1}^{N} \lambda_{s,k-j} A_{e,s} \| \leq \theta_2^{i-1} .
\]
(A36)

Moreover, from (A34), we have
\[
\| \tilde{t}_{k-i} \| = \| \Re v_{k-i} + \Psi \Phi w_{k-i} \| \leq \bar{\eta} ,
\]
(A37)

with \( \Re \triangleq - (\Psi \Phi G_1 M_1 T_1 + \Psi \Phi G_2 M_2 T_2 + L T_2) \). Then, from (A35)–(A37), we obtain (36) with \( \theta_2 \) and \( \bar{\eta} \) defined in (43). Furthermore, the result in (37) with \( \beta \) and \( \bar{\alpha} \) defined in (43), follows from applying Lemma 5, as well as triangle inequality, the facts that \( 0 \leq \hat{\lambda}_{i,k} \leq 1, \sum_{i=1}^{N} \hat{\lambda}_{i,k} = 1 \) and submultiplicativity of matrix norms.

Further, the steady state values are obtained by taking the limit from both sides of (38), (36), and (37), and assuming that \( \theta \triangleq \min(\theta_1, \theta_2) < 1 \).

\( \blacksquare \)

A.17 Proof of Corollary 4

Clearly \( \| A_{e,i} \| < 1 \) implies that \( \theta < 1 \), which is a sufficient condition for the convergence of errors by Theorem 4.

\( \blacksquare \)

A.18 Proof of Proposition 11

The stability of the observer and the noise attenuation level follow directly from Theorems 2 and 3, respectively. Moreover, we will show that the additional LMIs
\[
\kappa_1 I \leq P \leq \kappa_2 I ,
\]
(A38)
\[
(\kappa_1 \geq 1, \kappa_2 - \kappa_1 < 1) \lor (\kappa_2 \leq 1, \kappa_1 > 0.5) ,
\]
(A39)

imply that \( \theta_1 \triangleq \frac{\lambda_{\max}(P) - 1}{\lambda_{\min}(P)} < 1 \), which guarantees that \( \theta = \min(\theta_1, \theta_2) < 1 \) and therefore, the radii upper sequences are convergent by Theorem 4. To do so, we consider two cases:

• Case A) \( \kappa_1 \geq 1, \kappa_2 - \kappa_1 < 1 \). This, along with (A38) result in \( \lambda_{\max}(P) \geq \kappa_1 \geq 1 \) and \( |\lambda_{\max}(P) - 1| = \lambda_{\max}(P) - 1 \leq \kappa_2 - 1 < \kappa_1 \leq \lambda_{\min}(P) \). Hence, \( \theta_1 \triangleq \frac{|\lambda_{\max}(P) - 1|}{\lambda_{\min}(P)} < 1 \).

• Case B) \( \kappa_2 \leq 1, \kappa_1 > 0.5 \). This, in addition to (A38) return \( \lambda_{\max}(P) \leq \kappa_2 \leq 1 \) and \( \lambda_{\max}(P) - 1 = 1 - \lambda_{\max}(P) \leq 1 - \kappa_1 < \kappa_1 \leq \lambda_{\min}(P) \). So, \( \theta_1 \triangleq \frac{|\lambda_{\max}(P) - 1|}{\lambda_{\min}(P)} < 1 \).

\( \blacksquare \)