Author/s:
Yang, T;Murguia, C;Kuijper, M;Nešić, D

Title:
A multi-observer based estimation framework for nonlinear systems under sensor attacks

Date:
2020-09-01

Citation:
Yang, T., Murguia, C., Kuijper, M. & Nešić, D. (2020). A multi-observer based estimation framework for nonlinear systems under sensor attacks. Automatica, 119, https://doi.org/10.1016/j.automatica.2020.109043.

Persistent Link:
https://hdl.handle.net/11343/251857

A Multi-Observer Based Estimation Framework for Nonlinear Systems under Sensor Attacks

T. Yang [a], C. Murguia [a], M. Kuijper [a], D. Nešić [a]

Department of Electrical and Electronics Engineering, The University of Melbourne, Australia

Abstract

We address the problem of state estimation and attack isolation for general discrete-time nonlinear systems when sensors are corrupted by (potentially unbounded) attack signals. For a large class of nonlinear plants and observers, we provide a general estimation scheme, built around the idea of sensor redundancy and multi-observer, capable of reconstructing the system state in spite of sensor attacks and noise. This scheme has been proposed by others for linear systems/observers and here we propose a unifying framework for a much larger class of nonlinear systems/observers. Using the proposed estimator, we provide an isolation algorithm to pinpoint attacks on sensors during sliding time windows. Simulation results are presented to illustrate the performance of our tools.

Key words: Nonlinear observers; multi-observer; cyber-physical systems; sensor attacks.

1 Introduction

Networked Control Systems (NCSs) have emerged as a technology that combines control, communication, and computation, and offers the necessary flexibility to meet new demands in distributed and large scale systems. Recently, security of NCSs has become a very important issue as wireless communication networks increasingly serve as new access points for adversaries trying to disrupt the system dynamics. Cyber-physical attacks on control systems have caused substantial damage to a number of physical processes. A well-known example is the attack on Maroochy Shire Council's sewage control system in Queensland, Australia. The attacker hacked into the controllers that activate/deactivate valves causing a massive flooding to the surrounding areas. Another more recent incident is the StuxNet virus that targeted Siemens' supervisory control and data acquisition systems which are used in many industrial processes. These incidents show that strategic mechanisms to identify and deal with attacks on NCSs are needed.

In [4, 5, 7, 14, 19–23, 30, 31], a range of topics related to security of control systems have been discussed. In general, they provide analysis tools for quantifying the performance degradation induced by different classes of attacks and propose reaction and prevention strategies to counter their effect on the system dynamics. Most of the existing work, however, has considered control systems with linear dynamics, although in many engineering applications the dynamics of the plants being monitored and controlled is highly nonlinear. There are only a few results addressing the problem of state estimation under attacks for some classes of nonlinear systems. The recent work in [16] addresses the problem of sensor attack detection and state estimation for uniformly observable continuous-time nonlinear systems. For a class of power systems under sensor attacks, the authors in [9] provide an estimator of the system state using compressed sensing techniques. In [23], satisfiability modulo theory is used for state estimation for differentially flat systems with corrupted sensors. In our previous work [36, 37], the problem of state estimation and attack isolation for a class of nonlinear systems with *positive-slope nonlinearities* is considered. We provided an observer-based estimation/isolation strategy, using a bank of circle-criterion observers, which provides a robust estimate of the system state in spite of sensor attacks and effectively pinpoints attacked sensors.

The core of our estimation scheme is based on the work in [5], where the problem of state estimation for *continuous-time LTI systems* is addressed. The authors propose a multi-observer estimator, using a bank of Luenberger observers, which provides a robust estimate of the system

This work was supported by the Australian Research Council under the Discovery Project DP170104099.

Email addresses: tianciy@student.unimelb.edu.au (T. Yang), carlos.murguia@unimelb.edu.au (C. Murguia), mkuijper@.unimelb.edu.au (M. Kuijper), dnesic@.unimelb.edu.au (D. Nešić).

state in spite of sensor attacks. In this manuscript, we extend the results in [5,36,37] by considering systems with general nonlinear dynamics. We cast the multi-observer estimation scheme in terms of the existence of a bank of (local and practical) nonlinear observers with Input-to-State-Stable (ISS) (with respect to disturbances) estimator error dynamics. We consider the setting where the system has p sensors and up to $q < p$ of them are attacked. Following the multi-observer approach given in [5], we use a bank of observers to construct an estimator that provides a robust state estimate in the presence of false data injection attacks and noise.

The main idea behind the multi-observer estimator is the following: Each observer in the bank is driven by a different subset of sensors. Then, for every pair of observers in the bank, the estimator computes the difference between their estimates and selects the observers leading to the smallest difference. If there are attacks on some of the sensors, the observers driven by those sensors produce larger differences than the attack-free ones, in general, and thus they are not selected by the estimator. We first consider the noise-free case and show that our estimator converges to the true state of the system in spite of sensor attacks. Next, we consider the case when process disturbances and measurement noise are present. Assuming each observer's error is Input-to-State Stable (ISS) with respect to measurement noise and disturbances in the attack-free case, our estimator provides estimates whose errors satisfy an ISS-like property with respect to disturbances and independent of the attack signals. Compared to the estimation methods given in [9, 23], where no system disturbances and noise are considered, our estimation framework can deal with a much larger class of nonlinear systems at the price of having to design multiple observers. Finally, we provide an algorithm for isolating attacked sensors using the proposed estimator and assuming that upper bounds on the system noise are known. The idea behind our isolation algorithm is the following: For each pair of observers, when driven by attack-free sensors, the largest difference between their estimates is proved to be bounded by a threshold that depends on system noise bounds. For every time-step, we select and take the union of all the subsets of sensors such that the corresponding threshold is not crossed; then, the remaining sensors are isolated as attacked ones. To improve the isolation performance, we carry out the isolation over windows of N time-steps. That is, we select the subset of sensors that is isolated most often in every time window as the attacked ones. In [24, 29], the problem of isolation of attacked sensors for LTI systems is addressed using the *majority-vote* method and *satisfiability modulo theory*, respectively. Compared to those results, our isolation algorithm can be applied to nonlinear and noisy systems.

The remaining of the paper is organized as follows. Notation is given in Section 2. In Section 3, we present the multi-observer based estimator for the noise-free case. In Section 4, for the case with sensor noise and process disturbances, we prove that the observer-based estimator given in Section 3 provides ISS-like estimates of the system state (with respect to disturbances and noise) that are independent of sensor attacks. An algorithm for attack isolation is given in Section 5. Finally, we give concluding remarks in Section 6.

2 Notation

For any vector $v \in \mathbb{R}^n$, we denote v^J the stacking of all v_i, $i \in J$, $J \subset \{1,\dots,n\}$, $|v| = \sqrt{v^\top v}$, and the support set of v as $\text{supp}(v) = \{i \in \{1,\dots,n\} \,|v_i \neq 0\}$. For matrices $C \in \mathbb{R}^{p\times n}$, $C^\top = (c_1^\top,\dots,c_p^\top)$, we denote C^J the stacking of all rows $c_i \in \mathbb{R}^{1\times n}$, $i \in J$, $J \subset \{1,\dots,n\}$. For a sequence of vectors $\{v(k)\}_{k=0}^{\infty}$, $||v||_\infty := \sup_{k\geq 0} |v(k)|$. We say that a sequence $\{v(k)\}$ belongs to l_∞, $\{v(k)\} \in l_\infty$, if $||v||_\infty < \infty$. We denote uniformly distributed variables m in the interval (z_1, z_2) as $m \sim \mathcal{U}(z_1, z_2)$ and normally distributed with mean μ and variance σ^2 as $m \sim \mathcal{N}(\mu, \sigma^2)$. A continuous function $\alpha : [0, a) \to [0, \infty)$ is said to belong to class K, if it is strictly increasing and $\alpha(0) = 0$, [15]. Similarity, a continuous function $\beta : [0, a) \times [0, \infty) \to [0, \infty)$ is said to belong to class KL if, for fixed s, the mapping $\beta(r, s)$ belongs to class K with respect to r and, for fixed r, the mapping $\beta(r, s)$ is decreasing with respect to s and $\beta(r, s) \to 0$ as $s \to \infty$, [15].

3 Multi-Observer Estimator (Noise-free Case)

A multi-observer based estimator for continuous-time LTI systems has been proposed in [5]. Similarly, in [36], the authors give an estimator for nonlinear systems with positive-slope nonlinearities. Here, we generalize these results by considering general discrete-time nonlinear systems. Consider the nonlinear system

$$\begin{cases} x^+ = f(x,u), \\ y_i = h_i(x,u,a_i),\ i \in \{1,\dots,p\}, \end{cases} \tag{1}$$

with state $x \in \mathbb{R}^n$, input $u \in \mathbb{R}^{n_u}$, i-th sensor measurement $y_i \in \mathbb{R}$, stacked output $y := (y_1,\dots,y_p)^\top \in \mathbb{R}^p$, attack signal $a_i \in \mathbb{R}$, stacked attack vector $a := (a_1,\dots,a_p)^\top \in \mathbb{R}^p$, and functions $f : \mathbb{R}^n \times \mathbb{R}^{n_u} \to \mathbb{R}^n$ and $h_i : \mathbb{R}^n \times \mathbb{R}^{n_u} \times \mathbb{R} \to \mathbb{R}$. If the i-th sensor is not attacked, $a_i(k) = 0$ for $k \geq 0$; otherwise, sensor i is under attack and $a_i(k)$ is arbitrary and possibly unbounded. The *unknown* set of attacked sensors is denoted as W, $W \subset \{1,\dots,p\}$.

Assumption 1 *The set of attacked sensors does not change over time, i.e., W is constant (time-invariant) and $\text{supp}(a(k)) \subseteq W$, for all $k \geq 0$.*

Consider the observer

$$\begin{cases} z_J^+ = \Gamma_J(z_J, y^J, u), \\ \hat{x}_J = \eta_J(z_J, y^J, u), \end{cases} \tag{2}$$

Convergence	References
Global exponential	[11, 17, 18, 26, 32, 35, 36]
Global asymptotic	[3, 33, 34, 38]
Local exponential	[10, 18, 27, 28]
Local asymptotic	[2, 6, 8, 12, 25]
Finite-time	[13, 18]

Table 1
Systems/observers satisfying Definition 1 in the literature.

where $y^J \in \mathbb{R}^{\mathrm{card}(J)}$ denotes the stacking of all y_i, $i \in J$, $J \subset \{1, \ldots, n\}$, $z_J \in \mathbb{R}^{l_J}$ is the observer state, $\hat{x}_J \in \mathbb{R}^n$ denotes the estimate of the plant state, and $\Gamma_J : \mathbb{R}^{l_J} \times \mathbb{R}^{\mathrm{card}(J)} \times \mathbb{R}^{n_u} \to \mathbb{R}^{l_J}$ and $\eta_J : \mathbb{R}^{l_J} \times \mathbb{R}^{\mathrm{card}(J)} \times \mathbb{R}^{n_u} \to \mathbb{R}^n$ are some functions.

Definition 1 (Local Asymptotic Practical Observer). *System* (2) *is said to be a local asymptotic practical observer for system* (1) *if, for* $a^J(k) = 0$, $k \geq 0$, *there exists a set-valued map* $\mathcal{D}_J(x) \subseteq \mathbb{R}^{l_J}$, *such that, for any pair of initial conditions* $(x(0), z_J(0)) \in \mathbb{R}^n \times \mathcal{D}_J(x(0))$ *and* $e_J(k) := \hat{x}_J(k) - x(k)$, *there exist KL-function* $\beta_J(\cdot)$ *and* $\nu_J \geq 0$ *satisfying:* $|e_J(k)| \leq \beta_J(|e_J(0)|, k) + \nu_J$, $k \geq 0$.

In this manuscript, we assume that observers of form described in Definition 1 exist and are known for different subsets of sensors y^J, $J \subseteq \{1, \ldots, p\}$. Any technique available in literature can be used to construct these observers as long as the corresponding convergence properties satisfy Definition 1. Note that all observers guaranteeing global (local) asymptotic convergence satisfy Definition 1 with $\nu = 0$. In Table 1, we present a list of publications where design methods for nonlinear observers satisfying Definition 1 are given. We also list the corresponding convergence properties that these observers guarantee. The results in this paper apply to all the listed systems/observers.

Assumption 2 *At most* q *sensors are attacked, i.e.,*

$$\mathrm{card}(W) \leq q < \frac{p}{2}, \tag{3}$$

where q *denotes the largest integer such that for all* $J \subset \{1, \ldots, p\}$ *with* $\mathrm{card}(J) \geq p - 2q > 0$, *an observer of the form* (2) *exists for any* $y^J \in \mathbb{R}^{\mathrm{card}(J)}$.

Following the ideas in [5], we use a local asymptotic practical observer for each subset $J \subset \{1, \ldots, p\}$ of sensors with $\mathrm{card}(J) = p - q$ and for each subset $S \subset \{1, \ldots, p\}$ with $\mathrm{card}(S) = p - 2q$. By Assumption 2, among the p sensors, there exists at least one subset of sensors $\bar{I}$, $\bar{I} \subset \{1, \ldots, p\}$, with $\mathrm{card}(\bar{I}) = p - q$ satisfying $y^{\bar{I}} = h^{\bar{I}}(x, u)$, i.e., there is a set $\bar{I}$ of sensors that is attack-free and thus $a^{\bar{I}}(k) = 0$ for all $k \geq 0$. Then, in general, the difference between estimate $\hat{x}_{\bar{I}}(k)$ and the estimate $\hat{x}_S(k)$ given by any subset $S \subset \bar{I}$ with $\mathrm{card}(S) = p - 2q$ is smaller than the other subsets J with $\mathrm{card}(J) = p - q$ and $a^J(k) \neq 0$. This motivates the following estimation strategy.

For each subset $J \subset \{1, \ldots, p\}$ with $\mathrm{card}(J) = p - q$, define $\pi_J(k)$ as the largest deviation between the estimates $\hat{x}_J(k)$ and $\hat{x}_S(k)$ for any $S \subset J$ with $\mathrm{card}(S) = p - 2q$:

$$\pi_J(k) := \max_{S \subset J : \mathrm{card}(S) = p - 2q} |\hat{x}_J(k) - \hat{x}_S(k)|, \tag{4}$$

for all $k \geq 0$, and define the sequence $\sigma(k)$ as

$$\sigma(k) := \operatorname*{arg\,min}_{J \subset \{1,2,\ldots,p\} : \mathrm{card}(J) = p - q} \pi_J(k). \tag{5}$$

Then, as proven below, the estimate indexed by $\sigma(k)$:

$$\hat{x}(k) = \hat{x}_{\sigma(k)}(k), \tag{6}$$

is an asymptotic attack-free estimate of the system state. The following result uses the terminology presented above.

Theorem 1 *Consider system* (1), *observer* (2), *estimator* (4)-(6), *and the estimation error* $e(k) = \hat{x}_{\sigma(k)}(k) - x(k)$. *Let Assumption 1-2 be satisfied; then, there exist a constant* $\nu \geq 0$ *and a class* KL*-function* $\beta(\cdot)$ *satisfying:*

$$\begin{cases} |e(k)| \leq \beta(e_0, k) + \nu, \\ e_0 := \max\limits_{\substack{J \,:\, \mathrm{card}(J) \,=\, p \,-\, q \\ S \,:\, \mathrm{card}(S) \,=\, p \,-\, 2q}} \{|e_J(0)|, |e_S(0)|\}, \end{cases} \tag{7}$$

for all $k \geq 0$.

We omit the proof of Theorem 1 since we later prove a more general result in Section 4.

3.1 Application Examples

In this subsection, we show the performance of the proposed estimation scheme for two classes of nonlinear systems and observers.

High Gain Observers: Consider the nonlinear system

$$\begin{cases} x^+ = f(x), \\ y = h(x) + a, \end{cases} \tag{8}$$

with state $x \in \mathbb{R}^n$, output $y \in \mathbb{R}^p$, attack vector $a \in \mathbb{R}^p$, and functions $f : \mathbb{R}^n \to \mathbb{R}^n$ and $h : \mathbb{R}^n \to \mathbb{R}^p$.

Assumption 3 *The origin of* (8) *is locally stable* [15].

Consider the observer

$$\hat{x}_J^+ = f(\hat{x}_J) + K_J(y^J - h(\hat{x}_J)), \tag{9}$$

with state estimate $\hat{x}_J \in \mathbb{R}^n$ and observer gain matrix $K_J \in \mathbb{R}^{n \times \mathrm{card}(J)}$. The observer gain K_J is designed following the results in [27].

Proposition 1 *Let Assumption 3 be satisfied and* q *be the largest integer such that for all* $J \subset \{1, \ldots, p\}$ *with* $\mathrm{card}(J) \geq p - 2q$ *an observer of the form* (9) *for system* (8) *exists for any* $y^J \in \mathbb{R}^{\mathrm{card}(J)}$. *Then, for* $a^J(k) = 0$, $k \geq 0$, *there exists a set-valued map* $\mathcal{D}_J(x) \subsetneq \mathbb{R}^n$, *such that, for any* $(x(0), \hat{x}_J(0)) \in \mathbb{R}^n \times \mathcal{D}_J(x(0))$, *there are* $\lambda_J \in (0, 1)$ *and* $c_J > 0$ *satisfying* $|e_J(k)| \leq c_J \lambda_J^k |e_J(0)|$, $k \geq 0$, *where* $e_J = \hat{x}_J - x$.

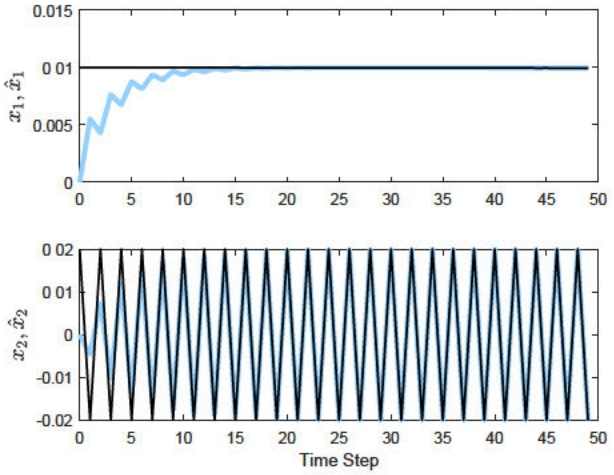

Fig. 1. Estimated states $\hat{x}$ converges to the true states x when $a_2 \sim \mathcal{U}(-10, 10)$. Legend: $\hat{x}$ (blue), true states (black).

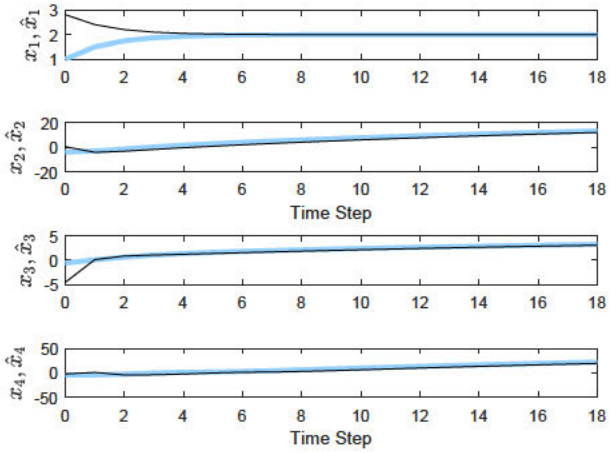

Fig. 2. Estimated states $\hat{x}$ converges to the true states x when $a_2 \sim \mathcal{U}(-10, 10)$. Legend: $\hat{x}$ (blue), true states (black)

Proof: Proposition 1 follows from [27, Theorem 3].

By Proposition 1, system (8) with observer (9) satisfy Definition 1 with $\beta(|e_J(0)|, k) = c_J \lambda_J^k |e_J(0)|$, $\nu_J = 0$, and some set-valued map $\mathcal{D}_J(x)$. Hence, we can write the following corollary of Theorem 1 and Proposition 1.

Corollary 1 *Consider system* (8)*, observer* (9)*, the estimator* (4)-(6)*, and the corresponding estimation error* $e(k) = \hat{x}_{\sigma(k)}(k) - x(k)$*. Let Assumptions 2 be satisfied; then, there exist* $c > 0$ *and* $\lambda \in (0, 1)$ *satisfying:* $|e(k)| \leq c\lambda^k e_0$*,* $k \geq 0$*, for* e_0 *as defined in* (7)*.*

Example 1: Consider the following nonlinear system subject to sensor attacks

$$\begin{cases} x_1^+ = x_1 - x_1^3 + x_2 x_1^2 - x_2^2 x_1^3, \\ x_2^+ = -x_2, \\ y_1 = 2x_1 + x_1^2, \\ y_2 = x_1 + x_2 + a_2, \\ y_3 = 2x_1 + x_2. \end{cases} \tag{10}$$

We have three sensors, i.e., $p = 3$. Using the design method given in [27], we have found that observers of the form (9) exist for each subset $J \subset \{1, 2, 3\}$ with $\text{card}(J) \geq 1$. By Assumption 2, $q = 1$, i.e., at most one sensor is attacked. We let $W = \{2\}$ and design an observer for each $J \subset \{1, 2, 3\}$ with $\text{card}(J) = 2$ and each $S \subset \{1, 2, 3\}$ with $\text{card}(S) = 1$. Therefore, totally $\binom{3}{2} + \binom{3}{1} = 6$ observers are designed. We fix the initial condition of the observers to $\hat{x}(0) = [0, 0]^\top$, select $(x_1(0), x_2(0)) \in \mathcal{N}(0, 1)$, and let $a_2 \sim \mathcal{U}(-10, 10)$. For $k \in [0, 49]$, we use (9),(4)-(6) to construct $\hat{x}(k)$. The performance of the estimator is shown in Figure 1.

Reduced Order Observers: Consider the system

$$\begin{cases} x^+ = Ax + f(x, y), \\ y = Cx + a, \end{cases} \tag{11}$$

with state $x \in \mathbb{R}^n$, output $y \in \mathbb{R}^p$, attack $a \in \mathbb{R}^p$, matrices $A \in \mathbb{R}^{n \times n}$ and $C \in \mathbb{R}^{p \times n}$, and nonlinear function $f : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^n$.

Assumption 4 *$f(x, y)$ is globally Lipschitz in x.*

Consider the partial output vector $y^J = C^J x + a^J$ and attack a^J, with $y^J, a^J \in \mathbb{R}^{\text{card}(J)}$, and the reduced state $\zeta_J = L_J x \in \mathbb{R}^{n - \text{card}(J)}$, where $L_J \in \mathbb{R}^{(n-\text{card}(J)) \times n}$ is such that $\begin{pmatrix} L_J^\top & (C^J)^\top \end{pmatrix}^\top$ is nonsingular. Let

$$(N_J, M_J) := \begin{pmatrix} L_J \\ C^J \end{pmatrix}^{-1};$$

then, $x = N_J \zeta_J + M_J y^J$, and we can write the dynamics of the reduced state ζ_J as

$$\zeta_J^+ = A_{L,J} \zeta_J + L_J \phi_J(\zeta_J, y^J) + B_{L,J} y^J, \tag{12}$$

where $A_{L,J} := L_J A N_J \in \mathbb{R}^{(n-\text{card}(J)) \times (n-\text{card}(J))}$, $B_{L,J} := L_J A M_J \in \mathbb{R}^{(n-\text{card}(J)) \times \text{card}(J)}$, and function $\phi_J(z_J, y^J) := f(N_J z_J + M_J y^J, y^J)$. Consider the reduced order observer

$$\begin{cases} z_J^+ = A_{LJ} z_J + \phi_J(z_J, y^J) + B_{LJ} y^J \\ \qquad\quad + K_J(y^{J+} - C^J \hat{x}_J^+), \\ \hat{x}_J = N_J z_J + M_J y^J, \end{cases} \tag{13}$$

with observer state $\hat{z}_J \in \mathbb{R}^{n - \text{card}(J)}$, estimated state $\hat{x}_J \in \mathbb{R}^n$, and observer matrix $K_J \in \mathbb{R}^{(n-\text{card}(J)) \times \text{card}(J)}$. We design K_J following the results in [38].

Proposition 2 *Let Assumption 4 be satisfied and q be the largest integer such that for all $J \subset \{1, \ldots, p\}$ with $\text{card}(J) \geq p - 2q$ an observer of the form* (13) *for system* (12) *exists for any $y^J \in \mathbb{R}^{\text{card}(J)}$. Then, for $a^J(k) = 0$, $k \geq 0$, and any $(x(0), z_J(0)) \in \mathbb{R}^n \times \mathbb{R}^{l_J}$, there exists a KL-function $\beta_J(\cdot)$ satisfying: $|e_J(k)| \leq \beta_J(|e_J(0)|, k)$, $k \geq 0$, where $e_J = \hat{x}_J - x$.*

Proof: Proposition 2 follows from [38, Theorem 4].

By Proposition 2, system (11) with observer (13) satisfy Definition 1 for some KL-function, $\nu_J = 0$, and set-valued map $D_J(x) = \mathbb{R}^n$. Hence, we can write the following corollary of Theorem 1 and Proposition 2.

Corollary 2 *Consider system* (11)*, observer* (13)*, the estimator* (4)-(6)*, and the corresponding estimation error* $e(k) = \hat{x}_{\sigma(k)}(k) - x(k)$*. Let Assumptions 2 be satisfied; then, there exists a class KL-function* $\beta(\cdot)$ *satisfying:* $|e(k)| \leq \beta(e_0, k)$*,* $k \geq 0$*, for* e_0 *as defined in* (7)*.*

Example 2: Consider the following nonlinear system under sensor attacks:

$$\begin{cases} x^+ = \begin{bmatrix} 0.5 & 0 & 0 & 0 \\ 0 & 0.8 & 1 & 0 \\ 0.5 & 0.1 & 0.3 & 0 \\ 0.3 & 1 & 0 & 0.5 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1.25 \tanh x_4 - 0.6 \end{bmatrix}, \\ y = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} x + \begin{bmatrix} 0 \\ a_2 \\ 0 \end{bmatrix}. \end{cases} \tag{14}$$

Using the design method proposed in [38], we have found that observers of the form (13) exist for each subset $J \subset \{1, 2, 3\}$ with $\text{card}(J) \geq 1$ and $p = 3$. By Assumption 2, $q = 1$, i.e., at most one sensor is attacked. For randomly selected initial conditions, we attack sensor two, i.e., $W = \{2\}$, and let $a_2 \sim \mathcal{U}(-10, 10)$. We use (13), (4)-(6) to reconstruct $x(k)$. The performance of the estimator is shown in Figure 2.

4 Robust Multi-Observer Based Estimator

The tools given in this section, generalize the results in [5, 36] by considering systems with general nonlinear dynamics, disturbances, and noise. Consider the system

$$\begin{aligned} x^+ &= F(x, u, d), \\ y_i &= g_i(x, u, m_i, a_i), i \in \{1, \ldots, p\}, \end{aligned} \tag{15}$$

with state $x \in \mathbb{R}^n$, input $u \in \mathbb{R}^{n_u}$, disturbance $d \in \mathbb{R}^s$, $\{d(k)\} \in l_\infty$, i-th sensor measurement $y_i \in \mathbb{R}$, stacked measurements $y := (y_1, \ldots, y_p)^\top \in \mathbb{R}^p$, attack signal $a_i \in \mathbb{R}$, measurement noise $m_i \in \mathbb{R}$, $\{m_i(k)\} \in l_\infty$, and nonlinear functions $F : \mathbb{R}^n \times \mathbb{R}^{n_u} \times \mathbb{R}^s \to \mathbb{R}^n$ and $g_i : \mathbb{R}^n \times \mathbb{R}^{n_u} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$.

Consider the observer

$$\begin{cases} z_J^+ = \Gamma_J(z_J, y^J, u), \\ \hat{x}_J = \eta_J(z_J, y^J, u), \end{cases} \tag{16}$$

where $z_J \in \mathbb{R}^{l_J}$ is the observer state, $\hat{x}_J \in \mathbb{R}^n$ denotes the state estimate, and $\Gamma_J : \mathbb{R}^{l_J} \times \mathbb{R}^{\text{card}(J)} \times \mathbb{R}^{n_u} \to \mathbb{R}^{l_J}$ and $\eta_J : \mathbb{R}^{l_J} \times \mathbb{R}^{\text{card}(J)} \times \mathbb{R}^{n_u} \to \mathbb{R}^n$ are some functions.

Convergence	References
Global exponential	[17, 26, 32, 36]
Global asymptotic	[1]

Table 2
Systems/observers satisfying Definition 2 in the literature.

Definition 2 (Local ISS Practical Observer). *System* (16) *is said to be a local asymptotic practical observer for system* (15) *if, for* $a^J(k) = 0$*,* $k \geq 0$*, there exists a set-valued map* $\mathcal{D}_J(x) \subseteq \mathbb{R}^{l_J}$*, such that for any pair of initial conditions* $(x(0), z_J(0)) \in \mathbb{R}^n \times \mathcal{D}_J(x(0))$ *and* $e_J = \hat{x}_J - x$*, there exist a* KL*-function* $\beta_J(\cdot)$*,* K*-functions* $\gamma_{1,J}(\cdot)$ *and* $\gamma_{2,J}(\cdot)$*, and constant* $\nu_J \geq 0$ *satisfying:*

$$\begin{aligned} |e_J(k)| \leq & \beta_J(|e_J(0)|, k) + \gamma_{1,J}(||m^J||_\infty) \\ & + \gamma_{2,J}(||d||_\infty) + \nu_J, \ k \geq 0. \end{aligned} \tag{17}$$

We assume that observers of form given in Definition 2 exist and are known for different subsets of sensors y^J, $J \subseteq \{1, \ldots, p\}$. In Table 2, we present a list of references where design methods for nonlinear observers satisfying Definition 2 can be found. All these observers can be used to construct the proposed estimator.

Assumption 5 *At most* q *sensors are attacked, i.e.,*

$$\text{card}(W) \leq q < \frac{p}{2}, \tag{18}$$

where q *denotes the largest integer such that for all* $J \subset \{1, \ldots, p\}$ *with* $\text{card}(J) \geq p - 2q > 0$*, an observer of the form* (16) *exists for any* $y^J \in \mathbb{R}^{\text{card}(J)}$*.*

Theorem 2 *Consider system* (15)*, observer* (16)*, estimator* (4)-(6)*, and the estimation error* $e(k) = \hat{x}_{\sigma(k)}(k) - x(k)$*. Let Assumptions 5 be satisfied; then, there exist a class* KL*-function* $\beta(\cdot)$*, class* K*-functions* $\gamma_1(\cdot)$ *and* $\gamma_2(\cdot)$*, and a constant* $\nu \geq 0$ *satisfying:*

$$\begin{cases} |e(k)| \leq \beta(e_0, k) + \gamma_1(||m||_\infty) + \gamma_2(||d||_\infty) + \nu, \\ e_0 := \max\limits_{\substack{J : \text{card}(J) = p - q \\ S : \text{card}(S) = p - 2q}} \{|e_J(0)|, |e_S(0)|\}. \end{cases} \tag{19}$$

for all $k \geq 0$ *and* $\{m(k)\}, \{d(k)\} \in l_\infty$*.*

Proof: Under Assumption 5, there exist at least one subset $\bar{I}$ with $\text{card}(\bar{I}) = p - q$ and $a^{\bar{I}}(k) = 0$ for all $k \geq 0$. Then, by definition 2, there exist a KL-function $\beta_{\bar{I}}(\cdot)$, class K-functions $\gamma_{1,\bar{I}}(\cdot)$ and $\gamma_{2,\bar{I}}(\cdot)$, and $\nu_{\bar{I}} \geq 0$ such that

$$|e_{\bar{I}}(k)| \leq \beta_{\bar{I}}(e_0, k) + \gamma_{1,\bar{I}}(||m^{\bar{I}}||_\infty) + \gamma_{2,\bar{I}}(||d||_\infty) + \nu_{\bar{I}}, \tag{20}$$

for all $k \geq 0$. For all $S \subset \bar{I}$ with $\text{card}(S) = p - 2q$, there exist a KL-function $\beta_S(\cdot)$, class K-functions $\gamma_{1,S}(\cdot)$ and $\gamma_{2,S}(\cdot)$, and $\nu_S \geq 0$ such that

$$|e_S(k)| \leq \beta_S(e_0, k) + \gamma_{1,S}(||m^S||_\infty) + \gamma_{2,S}(||d||_\infty) + \nu_S, \tag{21}$$

for all $k \geq 0$, which yields

$$\begin{aligned}\pi_{\bar{I}}(k) =& \max_{S\subset\bar{I}} |\hat{x}_{\bar{I}}(k) - \hat{x}_S(k)| \\ =& \max_{S\subset\bar{I}} |\hat{x}_{\bar{I}}(k) - x(k) + x(k) - \hat{x}_S(k)| \\ \leq& |e_{\bar{I}}(k)| + \max_{S\subset\bar{I}} |e_S(k)| \\ \leq& 2(\beta'(e_0,k) + \gamma'_1(||m^{\bar{I}}||_\infty) + \gamma'_2(||d||_\infty) + \nu'),\end{aligned} \tag{22}$$

for all $k \geq 0$, where

$$\gamma'_1(||m^{\bar{I}}||_\infty) = \max_{S\subset\bar{I}} \left\{\gamma_{1,\bar{I}}(||m^{\bar{I}}||_\infty), \gamma_{1,S}(||m^{\bar{I}}||_\infty)\right\},$$

$$\gamma'_2(||d||_\infty) = \max_{S\subset\bar{I}} \left\{\gamma_{2,\bar{I}}(||d||_\infty), \gamma_{2,S}(||d||_\infty)\right\}.$$

Under Assumption 5, for each subset $J \subset \{1,\dots,p\}$ with $\text{card}(J) = p-q$, there exists $\bar{S} \subset J$ with $\text{card}(\bar{S}) = p-2q$ such that $a^{\bar{S}}(k) = 0$ for all $k \geq 0$, and there exist a KL-function $\beta_{\bar{S}}(\cdot)$, class K-functions $\gamma_{1,\bar{S}}(\cdot)$ and $\gamma_{2,\bar{S}}(\cdot)$, and $\nu_{\bar{S}} \geq 0$ such that

$$|e_{\bar{S}(k)}| \leq \beta_{\bar{S}}(e_0,k) + \gamma_{1,\bar{S}}(||m^{\bar{S}}||_\infty) + \gamma_{2,\bar{S}}(||d||_\infty) + \nu_{\bar{S}}, \tag{23}$$

for all $k \geq 0$. From (4), by construction

$$\begin{aligned}\pi_{\sigma(k)}(k) =& \max_{S\supset\sigma(k):\text{card}(S)=2q} |\hat{x}_{\sigma(k)}(k) - \hat{x}_S(k)| \\ \geq& |\hat{x}_{\sigma(k)}(k) - \hat{x}_{\bar{S}}(k)|,\end{aligned}$$

using the above lower bound on $\pi_{\sigma(k)}(k)$ and the triangle inequality, we have that

$$\begin{aligned}|e_{\sigma(k)}(k)| =& |\hat{x}_{\sigma(k)}(k) - x(k)| \\ =& |\hat{x}_{\sigma(k)}(k) - \hat{x}_{\bar{S}}(k) + \hat{x}_{\bar{S}}(k) - x(k)| \\ \leq& |\hat{x}_{\sigma(k)}(k) - \hat{x}_{\bar{S}}(k)| + |e_{\bar{S}}(k)| \\ \leq& \pi_{\sigma(k)}(k) + |e_{\bar{S}}(k)| \\ \leq& \pi_{\bar{I}}(k) + |e_{\bar{S}}(k)|,\end{aligned} \tag{24}$$

for all $k \geq 0$. Hence, from (22) and (23), we have

$$|e_{\sigma(k)}| \leq 3(\beta_1(e_0,k) + \gamma_{1,1}(||m||_\infty) + \gamma_{2,1}(||d||_\infty) + \nu_1), \tag{25}$$

for all $k \geq 0$, where

$$\gamma_{1,1}(||m||_\infty) = \max\left\{\gamma'_1(||m||_\infty), \gamma_{1,\bar{S}}(||m||_\infty)\right\},$$

$$\gamma_{2,1}(||d||_\infty) = \max\left\{\gamma'_1(||d||_\infty), \gamma_{1,\bar{S}}(||d||_\infty)\right\}.$$

Inequality (25) is of the form (19) with KL-function $\beta(e_0,k) = 3\beta_1(e_0,k)$, nonnegative constant $\nu = 3\nu_1$, and K-functions $\gamma_1(||m||_\infty) = 3\gamma_{1,1}(||m||_\infty)$, and $\gamma_2(||d||_\infty) = 3\gamma_{2,1}(||d||_\infty)$. ■

4.1 Application Example

The following class of systems has been included in our preliminary work [36].

***Circle-Criterion Observers*:** Consider the system

$$\begin{cases} x^+ = Ax + Gf(Hx) + \rho(u,y), \\ y = Cx + a + m, \end{cases} \tag{26}$$

with state $x \in \mathbb{R}^n$, control $u \in \mathbb{R}^{n_u}$, output $y \in \mathbb{R}^p$, measurement noise $m \in \mathbb{R}^p$, $\{m(k)\} \in l_\infty$, and matrices $G \in \mathbb{R}^{n\times r}$ and $H \in \mathbb{R}^{r\times n}$. The term $\rho(u,y)$ is a known arbitrary real-valued vector that depends on the system inputs and outputs. The state-dependent nonlinearity $f(Hx)$ is an r-dimensional vector which each entry is a function of a linear combination of the states:

$$f_i = f_i\left(\sum_{j=1}^{n} H_{ij}x_j\right), \quad i = 1,\dots,r \tag{27}$$

where H_{ij} are the entries of matrix H.

Assumption 6 *For any $i \in \{1,\dots,r\}$:*

$$\frac{f_i(v_i) - f_i(w_i)}{v_i - w_i} \geq 0, \ \forall\ v_i, w_i \in \mathbb{R}, \ v_i \neq w_i. \tag{28}$$

Consider the circle-criterion observer

$$\begin{aligned}\hat{x}_J^+ =& A\hat{x}_J + Gf\big(H\hat{x}_J + K_J(C^J\hat{x}_J - y^J)\big) \\ &+ L_J(C^J\hat{x}_J - y^J) + \rho(u,y),\end{aligned} \tag{29}$$

with estimated state $\hat{x}_J \in \mathbb{R}^n$ and observer gain matrices $K_J \in \mathbb{R}^{r\times \text{card}(J)}$ and $L_J \in \mathbb{R}^{n\times\text{card}(J)}$. Matrices K_J and L_J are designed following the results in [36].

Proposition 3 *Let Assumption 6 be satisfied, and q be the largest integer such that for all $J \subset \{1,\dots,p\}$ with $\text{card}(J) \geq p - 2q > 0$ an observer of the form (29) for system (26) exists for any $y^J \in \mathbb{R}^{\text{card}(J)}$. Then, for $a^J(k) = 0$, $k \geq 0$, and any $(x(0), \hat{x}_J(0)) \in \mathbb{R}^n \times \mathbb{R}^n$, there exist $c_J > 0$, $\lambda_J \in (0,1)$, and $\gamma_{1,J} > 0$ satisfying: $|e_J(k)| \leq c_J\lambda_J^k|e_J(0)| + \gamma_{1,J}||m||_\infty$, $k \geq 0$, $\{m(k)\} \in l_\infty$, where $e_J = \hat{x}_J - x$.*

Proof: Proposition 3 follows from [36, Theorem 1].

By Proposition 3, system (26) with observer (29) satisfy Definition 2 with $\beta(|e_J(0)|,k) = c_J\lambda_J^k|e_J(0)|$, constant $d = 0$, linear function $\gamma_{1,J}$, $\nu_J = 0$, and set-valued map $D_J(x) = \mathbb{R}^n$. Hence, we can write the following corollary of Theorem 2 and Proposition 3.

Corollary 3 *Consider system (26), observer (29), the estimator (4)-(6), and the corresponding estimation error $e(k) = \hat{x}(k)_{\sigma(k)} - x(k)$. Let Assumptions 5 be satisfied; then, there exist $c > 0$, $\lambda \in (0,1)$, $\gamma_1 > 0$ satisfying: $|e(k)| \leq c\lambda^k e_0 + \gamma_1||m||_\infty$, $k \geq 0$, $\{m(k)\} \in l_\infty$, for e_0 as defined in (19).*

Example 3: Consider the following system subject to sensor noise and attacks

$$\begin{cases} x^+ = \begin{bmatrix} 1 & 0.1 \\ 0 & 1 \end{bmatrix} x + \begin{bmatrix} 0.05\sin(x_1+x_2) \\ 0.1\sin(x_1+x_2) \end{bmatrix}, \\ y = \begin{bmatrix} 3 & 3 & 6 & 1.2 & 1.5 \\ 0.3 & 0.6 & 0.9 & 12 & 15 \end{bmatrix}^\top x + m + a, \end{cases} \tag{30}$$

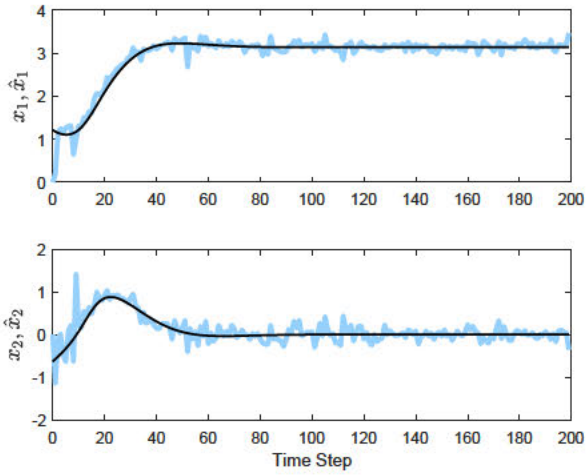

Fig. 3. Estimated states $\hat{x}$ converges to a neighbourhood of the true states x when $(a_2, a_5) \sim \mathcal{U}(-10, 10)$. Legend: $\hat{x}$ (blue), true states (black)

with $m_i \sim \mathcal{U}(-0.1, 0.1), i \in \{1, \ldots, 5\}$. Using the design method proposed in [36], we have found that observers of the form (29) exist for each subset $J \subset \{1, 2, 3, 4, 5\}$ with $\text{card}(J) \geq 1$ and $p = 5$. By Assumption 5, $q = 2$, i.e., at most two sensors are attacked. We design an observer for each $J \subset \{1, 2, 3, 4, 5\}$ with $\text{card}(J) = 3$ and each $S \subset \{1, 2, 3, 4, 5\}$ with $\text{card}(S) = 1$. Therefore, totally $\binom{5}{3} + \binom{5}{1} = 15$ observers are designed. We attack sensors two and five, i.e., $W = \{2, 5\}$, and let $(a_2, a_5) \sim \mathcal{U}(-10, 10)$. For $k \in [0, 199]$, we use (29),(4)-(6) to construct $\hat{x}(k)$. The performance of the estimator is shown in Figure 3.

5 Isolation of Attacked Sensors

Using the proposed estimation scheme, in our previous work [37], for a class of nonlinear systems with positive-slope nonlinearities, we have provided an algorithm for isolating sensor attacks. Here, we generalize this algorithm to deal with the larger class of systems (15). Consider system (15) and let q be the largest integer such that an observer of the form (16) satisfying Definition 2 exists for each subset $J \subset \{1, \ldots, p\}$ with $\text{card}(J) \geq p - 2q$.

Assumption 7 *Bounds on the process disturbance d and the sensor noise m are known, i.e.,*

$$||d||_\infty = \bar{d}, \ ||m||_\infty = \bar{m}, \tag{31}$$

where $\bar{d} \geq 0$ and $\bar{m} \geq 0$ are known constants.

To perform the isolation, we construct an observer satisfying Definition 2 for each subset $J \subset \{1, \ldots, p\}$ of sensors with $\text{card}(J) = p - q$ and each subset $S \subset \{1, \ldots, p\}$ with $\text{card}(S) = p - 2q$. Hence, by Definition 2, for $a^S(k) = 0$, $k \geq 0$, there exist a KL-function, $\beta_S(\cdot)$, K-functions, $\gamma_{1,S}(\cdot)$ and $\gamma_{2,S}(\cdot)$, and $\nu_S \geq 0$ satisfying:

$$|e_S(k)| \leq \beta_S(|e(0)|, k) + \gamma_{1,S}(\bar{m}) + \gamma_{2,S}(\bar{d}) + \nu_S, \tag{32}$$

for all $k \geq 0$. Note that, there always exist a k_S^* such that $\beta_S(|e(0)|, k) \leq \epsilon$, for any $\epsilon > 0$ and $k \geq k_S^*$. Then,

$$|e_S(k)| \leq \epsilon + \gamma_{1,S}(\bar{m}) + \gamma_{2,S}(\bar{d}) + \nu_S, \tag{33}$$

for all $k \geq k_S^*$. Define $\bar{k}^* := \max_{J,S} \{k_J^*, k_S^*\}$. By Assumption 5, there are at most q sensors under attack; then, we know there exists at least one $\bar{I} \subset \{1, \ldots, p\}$ with $\text{card}(\bar{I}) = p - q$ such that $a^{\bar{I}}(k) = 0, k \geq 0$, and

$$|e_{\bar{I}}(k)| \leq \epsilon + \gamma_{1,\bar{I}}(\bar{m}) + \gamma_{2,\bar{I}}(\bar{d}) + \nu_{\bar{I}}, \tag{34}$$

for all $k \geq k_{\bar{I}}^*$. Then, we have

$$\begin{aligned} \pi_{\bar{I}}(k) :=& \max_{S \subset \bar{I}} |\hat{x}_{\bar{I}}(k) - \hat{x}_S(k)| \\ =& \max_{S \subset \bar{I}} |\hat{x}_{\bar{I}}(k) - x(k) + x(k) - \hat{x}_S(k)| \\ \leq& |e_{\bar{I}}(k)| + \max_{S \subset \bar{I}} |e_S(k)|. \end{aligned} \tag{35}$$

From (33) and (34), we obtain

$$\pi_{\bar{I}}(k) \leq 2(\epsilon + \gamma'_{1,\bar{I}}(\bar{m}) + \gamma'_{2,\bar{I}}(\bar{d}) + \nu'_{\bar{I}}),$$

for all $k \geq \bar{k}^*$, where

$$\gamma'_{1,\bar{I}}(\bar{m}) := \max_{S \subset \bar{I}: \text{card}(S) = p - 2q} \left\{\gamma_{1,\bar{I}}(\bar{m}), \gamma_{1,S}(\bar{m})\right\},$$

and

$$\gamma'_{2,\bar{I}}(\bar{d}) := \max_{S \subset \bar{I}: \text{card}(S) = p - 2q} \left\{\gamma_{2,\bar{I}}(\bar{d}), \gamma_{2,S}(\bar{d})\right\}.$$

However, if the subset J of sensors is under attack at time k, i.e., $a^J(k) \neq 0$, then $\hat{x}_J(k)$ and $\hat{x}_S(k)$ in $\pi_J(k)$ are more inconsistent and produce larger $\pi_J(k)$. Define

$$\bar{\pi}_J := 2(\epsilon + \gamma'_{1J}(\bar{m}) + \gamma'_{2J}(\bar{d}) + \nu'_J), \tag{36}$$

for each $J \subset \{1, \ldots, p\}$ with $\text{card}(J) = p - q$, where

$$\gamma'_{1,J}(\bar{m}) := \max_{S \subset J: \text{card}(S) = p - 2q} \left\{\gamma_{1,J}(\bar{m}), \gamma_{1,S}(\bar{m})\right\},$$

and

$$\gamma'_{2,J}(\bar{d}) := \max_{S \subset J: \text{card}(S) = p - 2q} \left\{\gamma_{2,J}(\bar{d}), \gamma_{2,S}(\bar{d})\right\};$$

then, $\bar{\pi}_J$ can be used as a threshold to isolate attacked sensors. For all $k \geq \bar{k}^*$, we select from all the subsets $J \subset \{1, \ldots, p\}$ with $\text{card}(J) = p - q$, the ones that satisfy

$$\pi_J(k) \leq \bar{\pi}_J. \tag{37}$$

Denote as $\bar{W}(k)$ *the set of sensors that we regard as attack-free at time k*. We construct $\bar{W}(k)$ as the union of all subsets J satisfying (37):

$$\bar{W}(k) := \bigcup_{J \subset \{1, \ldots, p\}: \text{card}(J) = p - q, \pi_J(k) \leq \bar{\pi}_J} J. \tag{38}$$

Thus, the set $\{1, \ldots, p\} \setminus \bar{W}(k)$ is isolated as the set of attacked sensors at time k. Note, however, that, for small persistent attacks, it is still possible that for some $k \geq \bar{k}^*$ and some $J \subset \{1, \ldots, p\}$ with $\text{card}(J) = p - q$, $a^J(k) \neq 0$ but (37) still holds. This implies that $J \subset \bar{W}(k)$ even if

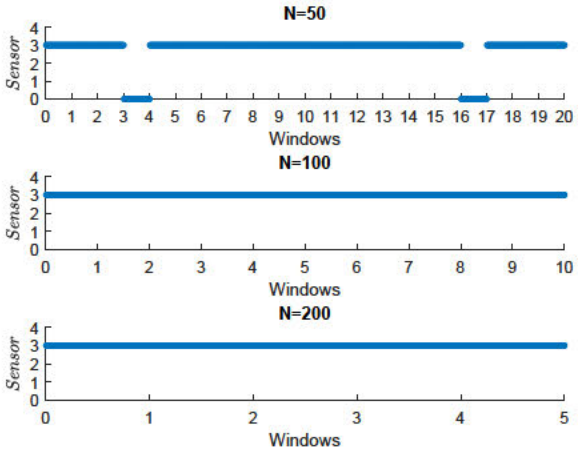

Fig. 4. Attack isolation, $a_3 \sim \mathcal{U}(-2, 2)$.

$a^J(k) \neq 0$ and would result in wrong isolation at time k. To improve the isolation performance, we carry out the isolation over windows of N time-steps, $N \in \mathbb{N}$. That is, for each $k \in [\bar{k}^* + (i-1)N, \bar{k}^* + iN]$, $i \in \mathbb{N}$, we compute and collect $\bar{W}(k)$ for every k in the window and select the subset J with $\text{card}(J) \geq p - q$ that is equal to $\bar{W}(k)$ most often in the i-th window. We denote this J as $J(i)$. Then, we select $\{1, \ldots, p\} \setminus J(i)$ as the set of sensors under attack in the i-th window. This isolation strategy is stated in Algorithm 1.

Example 4: Consider the nonlinear system subject to measurement noise and sensor attacks

$$\begin{cases} x^+ = \begin{bmatrix} 1 & 0.1 \\ 0 & 1 \end{bmatrix} x + \begin{bmatrix} 0.05 \sin(x_1 + x_2) \\ 0.1 \sin(x_1 + x_2) \end{bmatrix}, \\ y = \begin{bmatrix} 3 & 3 & 6 & 1.2 \\ 0.3 & 0.6 & 0.9 & 12 \end{bmatrix}^\top x + m + a, \end{cases} \tag{39}$$

with $m_i \sim \mathcal{U}(-0.5, 0.5)$ for $i \in \{1, 2, 3, 4\}$. Using the design method proposed in [36], we have found that circle-criterion observers of the form (29) satisfying Definition 2 exist for each subset $J \subset \{1, 2, 3, 4\}$ with $\text{card}(J) \geq 1$ and $p = 4$. It follows that, by Assumption 5, $q = 1$. We design a circle-criterion observer for each $J \subset \{1, 2, 3, 4\}$ with $\text{card}(J) = 3$ and each $S \subset \{1, 2, 3, 4\}$ with $\text{card}(S) = 2$. Therefore, in total, $\binom{4}{3} + \binom{4}{2} = 10$ observers are designed. We obtain their ISS gains by Monte Carlo simulations, initialize the observers at $\hat{x}(0) = x(0)$, select $(x_1(0), x_2(0))$ from a standard normal distribution, and fix $\epsilon = 0$. We let $N = 50, 100, 200$, and follow the evolution of Algorithm 1 for 1000 time-steps. We attack sensor three, i.e., $W = \{3\}$, and let $a_3 \sim \mathcal{U}(-2, 2)$. The isolation results are shown in Figures 4. In this figure, for visualization only, we depict $\tilde{A}_i = \emptyset$ (no isolated sensors) by sensor 0 being isolated in the i-th time window.

Algorithm 1 Attack Isolation.

1. Design an observer satisfying Definition 2 for each subset $J \subset \{1, \ldots, p\}$ with $\text{card}(J) = p - q$ and each subset $S \subset \{1, \ldots, p\}$ with $\text{card}(S) = p - 2q$.
2. Initialize the counter variable $n_J(i) = 0$ for all J with $\text{card}(J) \geq p - q$ and all $i \in \mathbb{Z}_{>0}$.
3. Compute $\bar{\pi}_J$ for each J with $\text{card}(J) = p - q$ as (5).
4. For $i \in \mathbb{Z}_{>0}$ and $\forall k \in [\bar{k}^* + (i-1)N, \bar{k}^* + iN - 1]$, compute $\pi_J(k)$, $\forall J$ with $\text{card}(J) = p - q$, as
$$\pi_J(k) = \max_{S \subset J : \text{card}(S) = p - 2q} |\hat{x}_J(k) - \hat{x}_S(k)|.$$
5. For all $k \in [\bar{k}^* + (i-1)N, \bar{k}^* + iN - 1]$, take the union of all the subsets J such that $\pi_J(k) \leq \bar{\pi}_J$:
$$\bar{W}(k) = \bigcup_{J \subset \{1, \ldots, p\} : \text{card}(J) = p - q, \pi_J(k) \leq \bar{\pi}_J} J.$$
6. For $k \in [\bar{k}^* + (i-1)N, \bar{k}^* + iN - 1]$, if $\bar{W}(k) = J$ for some J with $\text{card}(J) \geq p - q$, then update its corresponding counter variable as $n_J(i) = n_J(i) + 1$.
7. For all $i \in \mathbb{Z}_{>0}$, select the subset J with $\text{card}(J) \geq p - q$ that is equal to $\bar{W}(k)$ most often, i.e.,
$$J(i) = \arg\max_{J \in \{1, \ldots, p\} : \text{card}(J) \geq p - q} n_J(i).$$
8. For all $i \in \mathbb{Z}_{>0}$, the set of sensors potentially under attack is given by $\tilde{A}(i) = \{1, \ldots, p\} \setminus J(i)$.
9. For all $i \in \mathbb{Z}_{>0}$, return $\tilde{A}(i)$.

6 Conclusion

Following the idea of sensor redundancy and multi-observer in [5], a general estimation scheme has been proposed for a large class of nonlinear plants and observers, which provides robust estimate of the system state when a sufficiently small subset of sensors are corrupted by (potentially unbounded) attack signals and system plant as well as all sensors are affected by bounded noise. We have posed the multi-observer estimation scheme in terms of the existence of a bank of (local and practical) nonlinear observers with ISS (with respect to disturbances and noise) estimation error dynamics. We have proved that the proposed estimator provides ISS-like estimates of the system state with respect to disturbances only and independent of sensor attacks. This scheme has been proposed in [5], for linear systems/observers. Here, we have proposed a unifying framework for a much larger class of nonlinear systems/observers and provided the corresponding stability properties that the estimator yields in the nonlinear case. Using the proposed estimator, we have provided an isolation algorithm to pinpoint sensor attacks during finite time windows. Simulations results have been provided to illustrate the performance of our tools.

References

[1] M. Abbaszadeh and H. J. Marquez, "Robust H_∞ observer design for sampled-data Lipschitz nonlinear systems with exact and Euler approximate models," *Automatica*, vol. 44, no. 3, pp. 799–806, 2008.

[2] C. Califano, S. Monaco, and D. Normand-Cyrot, "On the observer design in discrete-time," *Systems and Control Letters*, vol. 49, no. 4, pp. 255–265, 2003.

[3] K. Chaib Draa, H. Voos, M. Alma, A. Zemouche, and M. Darouach, "An LMI-based H_∞ discrete-time nonlinear state observer design for an anaerobic digestion model," *20th IFAC World Congress*, 2017.

[4] M. S. Chong and M. Kuijper, "Characterising the vulnerability of linear control systems under sensor attacks using a system's security index," in *IEEE 55th Conference on Decision and Control (CDC)*, 2016, pp. 5906–5911.

[5] M. S. Chong, M. Wakaiki, and P. Hespanha, "Observability of linear systems under adversarial attacks *," *Proc. American Control Conf. (ACC)*, pp. 2439–2444, 2015.

[6] G. Ciccarella, M. D. Mora, and A. Germani, "A robust observer for discrete time nonlinear systems," *Systems and Control Letters*, vol. 24, no. 4, pp. 291–300, 1995.

[7] H. Fawzi, P. Tabuada, and S. Diggavi, "Secure estimation and control for cyber-physical systems under adversarial attacks," *IEEE Transactions on Automatic Control*, vol. 59, no. 6, pp. 1454–1467, 2014.

[8] A. Germani and C. Manes, "A discrete-time observer based on the polynomial approximation of the inverse observability map," *European journal of control*, vol. 15, no. 2, pp. 143–156, 2009.

[9] Q. Hu, D. Fooladivanda, Y. H. Chang, and C. J. Tomlin, "Secure state estimation and control for cyber security of the nonlinear power systems," *IEEE Transactions on Control of Network Systems*, pp. 1310 – 1321, 2017.

[10] S. Ibrir, "LPV approach to continuous and discrete nonlinear observer design," *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, no. 2, pp. 8206–8211, 2009.

[11] ——, "Circle-criterion approach to discrete-time nonlinear observer design," *Automatica*, vol. 43, no. 8, pp. 1432–1441, 2007.

[12] S. Ibrir, F. X. Wen, and C. Y. Su, "Observer-based control of discrete-time Lipschitzian non-linear systems: Application to one-link flexible joint robot," *International Journal of Control*, vol. 78, no. 6, pp. 385–395, 2005.

[13] T. Kaczorek, "Reduced-order perfect nonlinear observers of fractional descriptor discrete-time nonlinear systems," *International Journal of Applied Mathematics and Computer Science*, vol. 27, no. 2, pp. 245–251, 2017.

[14] S. H. Kafash, J. Giraldo, C. Murguia, A. A. Cardenas, and J. Ruths, "Constraining attacker capabilities through actuator saturation," in *proceedings of the American Control Conference (ACC)*, 2017.

[15] H. K. Khalil, *Nonlinear systems; 3rd ed.* Upper Saddle River, NJ: Prentice-Hall, 2002.

[16] J. Kim, C. Lee, H. Shim, Y. Eun, and J. H. Seo, "Detection of sensor attack and resilient state estimation for uniformly observable nonlinear systems," *IEEE 55th Conference on Decision and Control (CDC)*, pp. 1297–1302, 2016.

[17] G. Lu and D. Ho, "Observer design for a class of Lipschitz discrete-time systems," *IEEE International Conference on Control Applications*, pp. 1733–1738 Vol.2, 2004.

[18] P. Moraal and J. Grizzle, "Observer design for nonlinear systems with discrete-time measurements," *IEEE Transactions on Automatic Control*, vol. 40, no. 3, pp. 395–404, 1995.

[19] C. Murguia and J. Ruths, "Characterization of a CUSUM model-based sensor attack detector," in *IEEE 55th Conference on Decision and Control, CDC*, 2016, pp. 1303–1309.

[20] C. Murguia, N. van de Wouw, and J. Ruths, "Reachable sets of hidden CPS sensor attacks: analysis and synthesis tools," in *proceedings of the IFAC World Congress*, 2016, pp. 2088–2094.

[21] F. Pasqualetti, F. Dorfler, and F. Bullo, "Attack detection and identification in cyber-physical systems," *IEEE Transactions on Automatic Control*, vol. 58, pp. 2715–2729, 2013.

[22] Y. Shoukry, P. Nuzzo, A. Puggelli, A. Sangiovanni-Vincentelli, S.A.Seshia, and P. Tabuada, "Secure state estimation for cyber physical systems under sensor attacks: a Satisfiability Modulo Theory approach," *IEEE Transactions on Automatic Control*, vol. 62, no. 10, pp. 4917 – 4932, 2017.

[23] Y. Shoukry, P. Nuzzo, N. Bezzo, A. L. Sangiovanni-Vincentelli, S. A. Seshia, and P. Tabuada, "Secure state reconstruction in differentially flat systems under sensor attacks using satisfiability modulo theory solving," *54th IEEE Conference on Decision and Control (CDC)*, pp. 3804–3809, 2015.

[24] Y. Shoukry, A. Puggelli, P. Nuzzo, A. L. Sangiovanni-Vincentelli, S. A. Seshia, and P. Tabuada, "Sound and complete state estimation for linear dynamical systems under sensor attacks using Satisfiability Modulo Theory solving," *American Control Conference*, pp. 3818–3823, 2015.

[25] Y. Song and J. W. Grizzle, "The extended Kalman filter as a local asymptotic observer for discrete-time nonlinear systems," *J. Math. Syst. Estim. Control*, vol. 5, pp. 59–78, 1995.

[26] S. Sundaram, "State and unknown input observers for discrete-time nonlinear systems," in *IEEE 55th Conference on Decision and Control (CDC)*, 2016, pp. 7111–7116.

[27] V. Sundarapandian, "Observer design for discrete-time nonlinear systems," *Mathematical and computer modelling*, vol. 35, pp. 37–44, 2002.

[28] V. Sundarapandian and U. Pradesh, "General observers for discrete-time nonlinear systems," *Mathematical and Computer Modelling*, vol. 39, pp. 87–95, 2004.

[29] Z. Tang, M. Kuijper, M. S. Chong, I. Mareels, and C. Leckie, "Linear system security-detection and correction of adversarial sensor attacks in the noise-free case," *Automatica*, vol. 101, pp. 53–59, 2019.

[30] A. Teixeira, I. Shames, H. Sandberg, and K. H. Johansson, "Revealing stealthy attacks in control systems," *2012 50th Annual Allerton Conference on Communication, Control, and Computing, Allerton 2012*, pp. 1806–1813, 2012.

[31] K. G. Vamvoudakis, J. P. Hespanha, B. Sinopoli, and Y. Mo, "Detection in adversarial environments," *IEEE Transactions on Automatic Control*, vol. 59, no. 12, pp. 3209–3223, 2015.

[32] L. Xie, C. E. de Souza, and Y. Wang, "Robust filtering for a class of uncertain nonlinear systems: an H_∞ approach," in *Internaitional Journal of robust and nonlinear control*, vol. 6, no. 6, 1996, pp. 297–312.

[33] X. Xie, "Fuzzy observer of discrete-time nonlinear systems via an efficient maximum-priority-based switching mechanism," *ICCSS*, pp. 206–209, 2016.

[34] Y. Yalçin, "Discrete time immersion and invariance adaptive control via partial state feedback for systems in block strict feedback form," *European Journal of Control*, vol. 25, pp. 27–38, 2015.

[35] J. Yang, J. Back, J. H. Seo, and H. Shim, "Reduced-order dynamic observer error linearization," *IFAC Proceedings Volumes (IFAC-PapersOnline)*, pp. 915–920, 2010.

[36] T. Yang, C. Murguia, M. Kuijper, and D. Nešić, "A robust circle-criterion observer-based estimator for discrete-time nonlinear systems in the presence of sensor attacks," *IEEE 57th Conference on Decision and Control, CDC*, pp. 571–576, 2018.

[37] ——, "Attack detection and isolation for discrete-time nonlinear systems," *2018 Australian & New Zealand Control Conference (ANZCC)*, pp. 346–351, 2018.

[38] A. Zemouche and M. Boutayeb, "Observer design for Lipschitz nonlinear systems: the discrete-time case," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 53, no. 8, pp. 777–781, 2006.