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

Title:
An unknown input multiobserver approach for estimation and control under adversarial attacks

Date:
2021-03-01

Citation:
Yang, T., Murguia, C., Kuijper, M. & Nesic, D. (2021). An unknown input multiobserver approach for estimation and control under adversarial attacks. IEEE Transactions on Control of Network Systems, 8 (1), pp.475-486. https://doi.org/10.1109/TCNS.2020.3029160.

Persistent Link:
https://hdl.handle.net/11343/297877

An Unknown Input Multi-Observer Approach for Estimation and Control under Adversarial Attacks

Tianci Yang, Carlos Murguia, Margreta Kuijper, and Dragan Nešić

***Abstract*—We address the problem of state estimation, attack isolation, and control of discrete-time linear time-invariant systems under (potentially unbounded) actuator and sensor false data injection attacks. Using a bank of unknown input observers, each observer leading to an exponentially stable estimation error (in the attack-free case), we propose an observer-based estimator that provides exponential estimates of the system state in spite of actuator and sensor attacks. Exploiting sensor and actuator redundancy, the estimation scheme is guaranteed to work if a sufficiently small subset of sensors and actuators are under attack. Using the proposed estimator, we provide tools for reconstructing and isolating actuator and sensor attacks; and a control scheme capable of stabilizing the closed-loop dynamics by switching off isolated actuators. Simulation results are presented to illustrate the performance of our tools.**

***Index Terms*—Unknown input observers, cyber-physical systems, sensor and actuator attacks, linear systems, control.**

I. Introduction

Networked Control Systems (NCSs) have received considerable attention in recent years due to their numerous practical advantages (e.g., reduced weight, volume and installation costs, and better maintainability) when compared with traditional control systems where sensors and actuators communicate through point-to-point (wired) links. Modern critical engineering systems – energy, transportation, military, health care, and manufacturing – are being realized as large-scale NCSs. With the broad usage of NCSs, new security challenges have become an important issue as wireless communication networks increasingly serve as new access points for adversaries trying to disrupt the system. Cyber-physical attacks on NCSs have caused substantial damage to a number of engineering systems. A well-known example is the StuxNet virus that targeted Siemens's supervisory control and data acquisition systems. Another example is the false data injection attacks on power systems [1]. More recently, in 2014, the computers of a German steel mill were hacked causing the destruction of a blast furnace. These and many other recent incidents show that tools to identify and deal with attacks on NCSs are needed.

In [2]-[12], various security and privacy problems for linear control systems have been addressed and solved. In general, analysis (synthesis) tools are proposed to quantify (minimize) the performance degradation induced by different classes of attacks, e.g., bias injection, replay, zero dynamics, and denial-of-service. There are also some results addressing the nonlinear case. The problem of state estimation for nonlinear power systems under sensor attacks is solved in [13] using compressed sensing techniques. In [14], the authors address the problem of sensor attacks detection and state estimation for uniformly observable continuous-time nonlinear systems. In [15], Satisfiability Modulo Theory (SMT) tools are used for state estimation for nonlinear differentially flat systems with corrupted sensors. In our previous work [16], [17], the problem of state estimation and attack isolation for a class of nonlinear systems with *positive-slope nonlinearities* under sensor attacks is considered. Similar to the ideas given in [11], we provide an observer-based estimation/isolation strategy, using a bank of circle-criterion observers, which provides a robust state estimate in spite of sensor attacks. Most of the existing work assumes that actuators are attack-free and considers sensor attacks only. There are only a few results dealing with attacked actuators. For instance, in [18], the authors study the effect of actuator attacks on the performance of linear quadratic regulators. An adaptive control scheme that guarantees uniform ultimate boundedness of the closed-loop dynamics despite sensor and actuator attacks is given in [19]. In [20] and [21], the problem of state estimation under sensor and actuator attacks is addressed using compressed sensing and SMT-based techniques, respectively. Compared with the SMT-based estimator in [21], where state estimation is performed by analyzing sensor and actuator data collected within a window of finite length, observer-based estimators (the class of estimators considered here) show a higher promise of scalability since observers incorporate real-time sensor measurements as soon as they become available. This contrasts with the SMT scheme (and the compressed sensing technique in [20]) which requires memory for storing sensor/actuator information for a period of time. Moreover, the estimator in [20] (based on compressed sensing) heavily relies on a tuning parameter that controls the relative weight between sensor and actuator attacks. If this parameter is chosen incorrectly, the estimation algorithm leads to incorrect state reconstruction, and conditions for selecting it correctly are not provided. In contrast, we provide constructive techniques to design the estimator, and explicit sufficient conditions on the system dynamics, under which our estimator is guaranteed to work.

The core of our estimation scheme is inspired by the work in [11], where the problem of state estimation for *continuous-time linear time-invariant (LTI) systems* is addressed. The authors propose a multi-observer estimator, using a bank of Luenberger observers, that provides a robust estimate of the

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

The authors are with the Department of Electrical and Electronics Engineering, the University of Melbourne, Australia. tianciy@student.unimelb.edu.au

system state in spite of sensor attacks. The main idea behind their estimation scheme (and ours) is to place extra sensors (and actuators) in systems to create *redundancy*. Estimation under attacks would be impossible to achieve without redundancy (as many other authors have pointed out [20], [11], [22], [14], [12], [23]). Using redundant sensors/actuators is a very common technique in the research area of secure estimation and control, see, e.g., [20], [11], [22], [14], and references therein. Creating redundancy might be costly and lead to a conservative application of some of these estimation schemes, as it is not always possible to place extra sensors/actuators; however, when system security is critical, this is the price to pay. Some practical examples where sensor/actuator redundancy has been used to improve security can be found in [24], [25], [26], [27], [28], [29].

In this manuscript, using banks of Unknown Input Observers (UIOs) and exploiting redundancy in sensors and actuators, we address the problem of robust state estimation, attack reconstruction and isolation, and control for discrete-time LTI systems (with matrices (A, B, C)) under (potentially unbounded) actuator and sensor attacks. Unknown input observers are dynamical systems capable of estimating the state of the plant *without* using any input signals. If such an observer exists for the matrices $(A, B, \tilde{C}_i)$, where $\tilde{C}_i$ denotes a submatrix of C with fewer rows and the same number of columns, then, using banks of observers, we can perform state estimation and attack isolation when a sufficiently small subset of sensors are attacked (even if all inputs are under attack). The main idea behind our multi-observer estimator is the following. Each UIO in the bank is constructed using a triple $(A, B, \tilde{C}_i)$, i.e., the i-th observer is driven by the output signals associated with $\tilde{C}_i$ only. If the outputs corresponding to $\tilde{C}_i$ are attack-free, this UIO produces an exponentially stable estimation error. For every pair of UIOs in the bank, we compute the largest difference between their estimates. Then, we select the pair leading to the smallest difference and prove that these observers reconstruct the state of the system exponentially. If a UIO does not exist for $(A, B, \tilde{C}_i)$, but it does for $(A, \tilde{B}_i, \tilde{C}_i)$, where $\tilde{B}_i$ is a submatrix of B with fewer columns and the same number of rows, i.e., the i-th observer does *not* use the input signals associated with $\tilde{B}_i$, but it does use the remaining input signals and the output signals corresponding to $\tilde{C}_i$, then using banks of these UIOs, we can use similar ideas to perform state estimation and attack isolation at the price of only being able to isolate when a sufficiently small subset of actuators and sensors are under attack. If the inputs corresponding to $\tilde{B}_i$ include all the attacked ones and the outputs corresponding to $\tilde{C}_i$ are attack-free, this UIO produces exponentially stable estimation error. For every pair of UIOs in the bank, we compute the largest difference between their estimates and select the pair leading to the smallest difference. We prove that these observers provide exponential estimate of the system state. Once we have an estimate of the state, we provide tools for reconstructing attack signals using model matching techniques. Attacked actuators and sensors are isolated by simply checking the sparsity of the estimated attack signals. Finally, after obtaining state estimates and isolation has been performed, we provide a control scheme for stabilizing the closed-loop dynamics. In the case with sensor attacks only (no actuators attacks), we show that a separation principle between estimation and control holds and the system can be stabilized by closing the loop with the multi-observer estimator and a static output feedback controller. When both sensors and actuators are attacked, we propose an effective technique to stabilize the system by switching off the isolated actuators, and closing the loop with a multi-observer based output time-varying feedback controller. Because attack signals might be zero for some time instants, actuators isolated as attack-free might arbitrarily switch among all the supersets of the set of attack-free actuators. Therefore, we need a controller able to stabilize the closed-loop dynamics under the arbitrary switching induced by turning off the isolated actuators. To achieve this, we assume that a *state* feedback controller that stabilizes the switching closed-loop system exists, and use this controller together with the multi-observer estimator to stabilize the system. We use Input-to-State Stability (ISS) [30] of the closed-loop system with respect to the exponentially stable estimation error to conclude on stability of the closed-loop dynamics. Compared to the adaptive controller proposed in [19], where a particular class of attacks is considered and ultimate boundedness of the closed-loop system is guaranteed only, our controller is able to drive the system state asymptotically to the origin under arbitrary and potentially unbounded attack signals.

The paper is organized as follows. In Section II, we present some preliminary results needed for the subsequent sections. In Section III, we introduce the proposed UIO-based estimation schemes. In Section IV, a method for isolating attacked sensors and actuators is described. The proposed control scheme is given in Section V. Finally, in Section VI, we give concluding remarks.

II. Preliminaries

A. Notation

We denote the set of real numbers by $\mathbb{R}$, the set of natural numbers by $\mathbb{N}$, the set of integers by $\mathbb{Z}$, and $\mathbb{R}^{n\times m}$ the set of $n \times m$ matrices for any $m, n \in \mathbb{N}$. 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 $\text{supp}(v) = \{i \in \{1, \dots, n\} \,|v_i \neq 0\}$. 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 $c_i \in \mathbb{R}^{1\times n}$, $i \in J$, $J \subset \{1, \dots, n\}$. Set J is called a superset of set S if $S \subseteq J$. We denote the cardinality of a set S as $\text{card}(S)$. The binomial coefficient is denoted as $\binom{a}{b}$, where a, b are nonnegative integers. We denote a variable m uniformly distributed 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 $m \sim \mathcal{N}(\mu, \sigma^2)$. The notation $\mathbf{0}_n$ and I_n denote the zero matrix and the identity matrix of dimension $n\times n$, respectively. We simply write $\mathbf{0}$ and I when their dimensions are evident. For discrete-time signals, $x(k)$, at time index $k \in \mathbb{N}$, we often omit the index k and simply denote $x(k)$ as x. Similarly, we often write $x(k + 1)$ as x^+. Finally, for square matrices $A \in \mathbb{R}^{n\times n}$, we call matrix A *stable* if all the eigenvalues of A are inside the unit circle on the complex plane.

III. Estimation

In [11], the problem of state estimation for continuous-time LTI system under sensor attacks is solved using a bank of Luenberger observers. Inspired by these results, we use banks of UIOs to estimate the state of the system when sensor and actuator attacks both occur. We consider discrete-time linear time-invariant systems under sensor and actuator attacks:

$$\begin{cases} x^+ = Ax + B(u + a_u), \\ y = Cx + a_y, \end{cases} \tag{1}$$

with state $x = x(k) \in \mathbb{R}^n$, at time-step $k \in \mathbb{N}$, state at the next step $x^+ = x(k+1) \in \mathbb{R}^n$, output $y(k) \in \mathbb{R}^{n_y}$, and input $u(k) \in \mathbb{R}^{n_u}$. Vectors $a_u(k) = (a_{u1}, \ldots, a_{un_u})(k)^\top \in \mathbb{R}^{n_u}$ and $a_y(k) = (a_{y1}, \ldots, a_{yn_y})(k)^\top \in \mathbb{R}^{n_y}$ denote, respectively, sensor and actuator attacks at time-step $k \in \mathbb{N}$. The i-th entry of $a_y(k)$ ($a_u(k)$) is zero for $k \geq 0$, $a_{yi}(k) = 0$ ($a_{ui}(k) = 0$), if the i-th sensor (actuator) is attack-free; otherwise, $a_{yi}(k) \neq 0$ ($a_{ui}(k) \neq 0$) for some $k \in \mathcal{K} \subseteq \mathbb{N}$ and can be arbitrarily large. Let $W_u \subset \{1, \ldots, n_u\}$ and $W_y \subset \{1, \ldots, n_y\}$ denote the *unknown* set of attacked actuators and sensors, respectively. Matrices A, B, C are of appropriate dimensions, and B has full column rank.

We assume the adversary attacks a sensor (actuator) by modifying sensor measurements (input signals) to any arbitrary value. This attack is modeled as an additive term in (1) without any restriction (statistical or otherwise). From a practical point of view, an attack on a sensor could either be interpreted as an attack on the node, or an attack on the communication link between the sensor and the receiver device. Similarly, an attack on an actuator could either be interpreted as an attack on the node, or an attack on the communication link from the controller to the actuator.

Assumption 1 *The sets of attacked actuators and sensors do not change over time, i.e., sets* $W_u \subseteq \{1, \ldots, n_u\}$ *and* $W_y \subset \{1, \ldots, n_y\}$ *are time-invariant and* $\mathrm{supp}(a_u(k)) \subseteq W_u$ *and* $\mathrm{supp}(a_y(k)) \subseteq W_y$*, for all* $k \geq 0$*.*

A. Completely Unknown Input Observers

We first treat $(u + a_u)$ as an unknown input to system (1). For a subset J_s of sensors, consider a UIO of the form:

$$\begin{cases} z_{J_s}^+ = N_{J_s} z_{J_s} + L_{J_s} y^{J_s}, \\ \hat{x}_{J_s} = z_{J_s} + E_{J_s} y^{J_s}, \end{cases} \tag{2}$$

where $z_{J_s} \in \mathbb{R}^n$ is the state of the observer, $z_{J_s}^+ \in \mathbb{R}^n$ represents the state of the observer at the next time step, $\hat{x}_{J_s} \in \mathbb{R}^n$ denotes the estimate of the system state, and $(N_{J_s}, L_{J_s}, E_{J_s})$ are observer matrices of appropriate dimensions to be designed. In [31], it is verified that if $(N_{J_s}, L_{J_s}, E_{J_s})$ satisfy the following equations:

$$\begin{cases} N_{J_s}(I - E_{J_s} C^{J_s}) + L_{J_s} C^{J_s} + (E_{J_s} C^{J_s} - I)A = \mathbf{0}, \\ (E_{J_s} C^{J_s} - I)B = \mathbf{0}; \end{cases} \tag{3}$$

then, the estimation error $e_{J_s} = \hat{x}_{J_s} - x$ satisfies:

$$e_{J_s}^+ = N_{J_s} e_{J_s}. \tag{4}$$

If N_{J_s} is stable, system (2) is called a UIO for (1). In [31], it is proved that such an observer exists if and only if the following two conditions are satisfied:

($\mathbf{c_1}$) $\mathrm{rank}(C^{J_s} B) = \mathrm{rank}(B) = n_u$.

($\mathbf{c_2}$) The pair $(C^{J_s}, A - E_{J_s} C^{J_s} A)$ is detectable.

Let q be the largest integer such that, for all subsets $J_s \subset \{1, \ldots, n_y\}$ with $\mathrm{card}(J_s) \geq n_y - 2q > 0$, conditions (c_1) and (c_2) are satisfied. Then, observer (2) can be constructed for any C^{J_s} with $\mathrm{card}(J_s) \geq n_y - 2q$ by solving (3) for a stable matrix N_{J_s}. Hence, for such an observer, if $a_y^{J_s}(k) = 0$ for $k \geq 0$, there exist $c_{J_s} > 0$ and $\lambda_{J_s} \in (0, 1)$ satisfying: $|e_{J_s}(k)| \leq c_{J_s} \lambda_{J_s}^k |e_{J_s}(0)|$, for all $e_{J_s}(0) \in \mathbb{R}^n$, $k \geq 0$, where $e_{J_s} = \hat{x}_{J_s} - x$. See [31] for details.

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

$$\mathrm{card}(W_y) \leq q < \frac{n_y}{2}, \tag{5}$$

where q *is the largest positive integer satisfying* (c_1) *and* (c_2)*.*

Lemma 1 *Under Assumption 2, among each set of* $n_y - q$ *sensors, at least* $n_y - 2q > 0$ *of them are attack-free.*

Proof: Lemma 1 follows trivially from Assumption 2. ■

Let Assumption 2 be satisfied. Inspired by the ideas in [11], we use a UIO for each subset $J_s \subset \{1, \ldots, n_y\}$ of sensors with $\mathrm{card}(J_s) = n_y - q$ and for each subset $S_s \subset \{1, \ldots, n_y\}$ of sensors with $\mathrm{card}(S_s) = n_y - 2q$. Under Assumption 2, there exists at least one set $\bar{J}_s \subset \{1, \ldots, n_y\}$ with $\mathrm{card}(\bar{J}_s) = n_y - q$ such that $a_y^{\bar{J}_s}(k) = 0$ for all $k \geq 0$. Then, the estimate given by the UIO for $\bar{J}_s$ is a correct estimate, and the estimate given by the UIO for any $S_s \subset \bar{J}_s$ with $\mathrm{card}(S_s) = n_y - 2q$ is consistent with that given by $\bar{J}_s$. This motivates the following estimation strategy.

For each set J_s with $\mathrm{card}(J_s) = n_y - q$ and $k \geq 0$, we define $\pi_{J_s}(k)$ as the largest deviation between $\hat{x}_{J_s}(k)$ and $\hat{x}_{S_s}(k)$ that is given by any $S_s \subset J_s$ with $\mathrm{card}(S_s) = n_y - 2q$, i.e.,

$$\pi_{J_s}(k) := \max_{S_s \subset J_s : \mathrm{card}(S_s) = n_y - 2q} |\hat{x}_{J_s}(k) - \hat{x}_{S_s}(k)|, \tag{6}$$

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

$$\sigma_s(k) := \operatorname*{arg\,min}_{J_s \subset \{1, \ldots, n_y\} : \mathrm{card}(J_s) = n_y - q} \pi_{J_s}(k). \tag{7}$$

Then, as proved below, the estimate indexed by $\sigma_s(k)$:

$$\hat{x}(k) := \hat{x}_{\sigma_s(k)}(k), \tag{8}$$

is an exponential attack-free estimate of the system state. For simplicity and without loss of generality, for all J_s and S_s, $z_{J_s}(0)$ and $z_{S_s}(0)$ are chosen so that $\hat{x}_{J_s}(0) = \hat{x}_{S_s}(0) = \hat{x}(0)$. The following result summarizes the ideas presented above.

Theorem 1 *Consider system* (1)*, observer* (2)*, and the completely unknown multi-observer estimator* (6)-(8)*. Define the estimation error* $e(k) := \hat{x}_{\sigma_s(k)}(k) - x(k)$*, and let conditions* (c_1)*-*(c_2) *and Assumptions 1-2 be satisfied; then, there exist constants* $\bar{c} > 0$ *and* $\bar{\lambda} \in (0, 1)$ *satisfying:*

$$|e(k)| \leq \bar{c} \bar{\lambda}^k |e(0)|, \tag{9}$$

for all $e(0) \in \mathbb{R}^n$*,* $k \geq 0$*.*

Proof: See Appendix A. ■

Remark 1 *Together, Assumptions 1 and 2, are sufficient conditions for the problem to have a solution in our setting. They encompass the existence of the observers and the maximum number of sensors that can be attacked for the multi-observer estimator to work. On the other hand, the condition that no more than half of sensors are attacked is a necessary condition for any secure estimation scheme to exist. This is an assumption made in all related work, see, e.g.,* [20], [11], [22], [14], [27], [32], [33], [34], [12], [35], [23].

Remark 2 *Regarding computational complexity of the scheme, each observer* (2) *needs to perform a matrix multiplication of a* $2n \times (n+m)$ *matrix (where* n *is the dimension of the system,* $m = \text{card}(J_s) \leq n_y$*, and* n_y *is the number of sensors) and an* $(n+m)$*-dimensional vector. Then, at every time-step, the computational complexity is upper bounded by* $O(n^2+n_y)$ [36]*. This is for one observer, whereas we have* $\binom{n_y}{n_y-q}$ *observers. So at every time-step, the complexity is of order* $O((n^2+n_y)\binom{n_y}{n_y-q})$*. It grows exponentially with the number of sensors. We acknowledge that this is a limitation for applications with hundreds of sensors, but for medium-size systems with tens of sensors, the computational complexity is manageable by current computers. Also, in many practical applications, e.g., vehicle dynamics, water networks, traffic monitoring systems, and batch reactors (where the time-scale of the dynamics is slow in comparison with computation/communication time), computation time is not a practical concern anymore since computations are performed much faster than the evolution of the system dynamics. We remark that most of the related work,* [11], [14], [22]*, suffers from the same high complexity issue as the problem itself is combinatorial.*

Remark 3 *When system* (1) *is perturbed by bounded process and measurement disturbances, the proposed estimator* (6)-(8) *provides Input-to-State Stable (ISS) estimates of the system state (with respect disturbances) if Assumptions 1 and 2 are satisfied. This can be rigorously proved since each UIO in the bank is ISS with respect to disturbances in the absence of attacks.*

Remark 4 *The underlying mechanism behind our estimator is sensor and actuator redundancy. The idea is to place extra sensors/actuators in systems to create this redundancy. Indeed, this might be costly and lead to a conservative application of the scheme, as it would not be always possible to place extra different sensors/actuators. Note, however, that placing extra different sensors is not the only alternative. First, different sensors (i.e., strictly different rows in* C*) is not necessarily required. We could have as many repeated rows as needed – meaning that we could have repeated sensors multiple times. This is important as often in practice one only finds standard commercial sensors and thus having sensors measuring arbitrary combinations of states is not possible. Another option could be to create virtual sensors out of actual physical ones. For instance, if we only have two physical sensors but these two sensors render the system observable, we could artificially create extra virtual sensors by computing linear combinations of them. That is, for two scalar sensors, say* y_1 *and* y_2*, the set* $\{\alpha_1 y_1 + \alpha_2 y_2 | \alpha_1, \alpha_2 \in \mathbb{R}\}$ *characterizes an infinite set of potential virtual sensors computed out of* y_1 *and* y_2*. Hence, instead of buying costly extra sensors, we could compute virtual ones before transmission, and send both, actual and virtual sensor measurements, through the unsecured channel.*

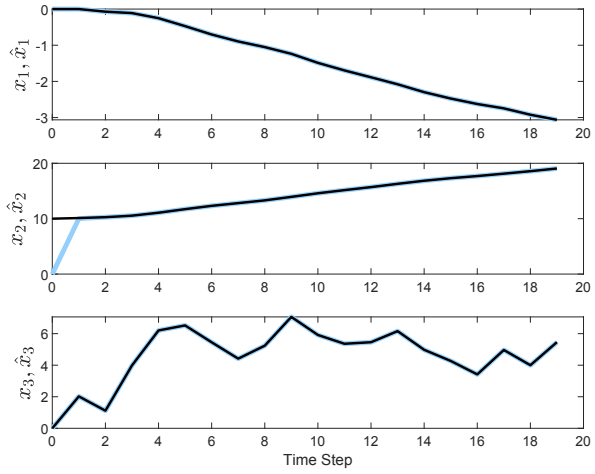

Fig. 1. Estimated states $\hat{x}$ converges to the true states x when $a_u, a_{y2} \sim \mathcal{U}(-10, 10)$. Legend: $\hat{x}$ (blue), true states (black)

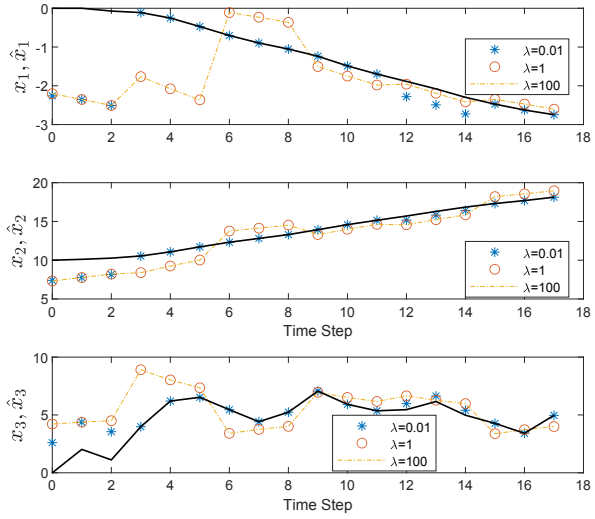

Fig. 2. Performance of estimator [20] for various values of λ, $a_u, a_{y2} \sim \mathcal{U}(-10, 10)$.

Example 1: Consider the following third-order continuous-time longitudinal vehicle dynamics:

$$\begin{bmatrix} \dot{p} \\ \dot{v} \\ \dot{\delta} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -\frac{1}{\tau} \end{bmatrix} \begin{bmatrix} p \\ v \\ \delta \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \frac{1}{\tau} \end{bmatrix} u, \tag{10}$$

where $p \in \mathbb{R}$, $v \in \mathbb{R}$, and $\delta \in \mathbb{R}$ denote, respectively, position, velocity, and acceleration of the vehicle, and $\tau \in \mathbb{R}_{>0}$ is the inertia time-lag in the power train. This model has been extensively used in vehicle platooning research, see, e.g., [37]-[39], and references therein. The vehicle is equipped with a Light Detection and Ranging (LIDAR) sensor and a Inertial Measurement Unit (IMU) that, combined, provide measurements of p, v, and δ. It can be verified that using (p, v, δ) as output (i.e., $C = I_3$) leads to no solution of the observer equations in (3) for any submatrix C^{J_s}. Thus, we cannot use the proposed estimator with this C. However, we could construct redundant virtual sensors out of (p, v, δ), and send these virtual measurements through the unsecured network instead (see Remark 4 for details). For this example, we use the following virtual sensors for estimation:

$$y = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} p \\ v \\ \delta \end{bmatrix} =: Cx. \tag{11}$$

We let $\tau = 0.4$, and exactly discretize (10) at the sampling time-instants (assuming a zero-order hold) with sampling interval of 0.1 seconds. The resulting discrete-time system under attacks is of the form (1), with sensor matrix C and state x as defined above, and (A, B) given by

$$A = \begin{bmatrix} 1 & 0 & -0.0354 \\ 0 & 1 & 0.0885 \\ 0 & 0 & 0.7788 \end{bmatrix}, \; B = \begin{bmatrix} 0 \\ 0.0125 \\ 0.2188 \end{bmatrix}. \tag{12}$$

It can be verified that a UIO of the form (2) exists for every two sensors; then, $4-2q = 2$, i.e., $q = 1$, and thus at most one sensor could be attacked. We let the actuator and the second sensor be attacked. We select $u = 5$, $a_u, a_{y2} \sim \mathcal{U}(-10, 10)$, and $x(0) = [0, 10, 0]^\top$. We design a UIO for each J_s with $\mathrm{card}(J_s) = 3$, and for each S_s with $\mathrm{card}(S_s) = 2$. Therefore, totally $\binom{4}{3} + \binom{4}{2} = 10$ UIOs are designed. They are all initialized at $\hat{x}(0) = [0, 0, 0]^\top$. For $k \in [0, 19]$, the estimator (2), (6)-(8) is used to construct $\hat{x}(k)$. The performance of the multi-observer estimator is shown in Figure 1. For comparison, the estimator in [20, Eq. (12)] is used to construct state estimates also. The performance is shown in Figure 2. The execution time of the proposed multi-observer estimator to reconstruct $\hat{x}(k)$, $k \in [0, 19]$, is 0.007 seconds. For comparison, the estimator in [20, Eq. (12)] takes 6.963 seconds to compute the estimate for $k \in [0, 17]$. The execution time of our scheme is mainly due to the iteration of observers in the bank (for given observer gains) and the procedure (6)-(8) for selecting the best estimate. For comparison, we run the estimation algorithm in [20], which requires solving an l_1-norm optimization problem (see [20, Eq. (12)]).

Remark 5 *As stated by the authors in [20], when sensor and actuator attacks occur at the same time, the performance of the estimator in [20] (based on compressed sensing techniques) heavily relies on a tuning parameter, say $\lambda \in \mathbb{R}$, that controls the relative weight between sensor and actuator attacks. If this parameter is chosen incorrectly, these estimation algorithms lead to incorrect state reconstruction, and conditions for selecting it correctly are not provided (known). Figure 2 shows state estimates for $\lambda = 0.01, 1, 100$. Note that we have a good estimate for $\lambda = 0.01$ and poor estimates for $\lambda = 1, 100$. In contrast, we provide constructive techniques to design the multi-observer based estimator, and explicit sufficient conditions on the system dynamics, which, if satisfied, our estimator is guaranteed to work.*

B. Partially Unknown Input Observers

Here, we are implicitly assuming that either condition (c_1) or (c_2) (or both) cannot be satisfied for any C^{J_s} with $\mathrm{card}(J_s) = n_y - 2q$ with $q \geq 1$. Let B be partitioned as $B = [b_1, \ldots, b_i, \ldots, b_{n_u}]$ where $b_i \in \mathbb{R}^{n \times 1}$ is the i-th column of B. Then, the attacked system (1) can be written as

$$\begin{cases} x^+ = Ax + Bu + b_{W_u} a_u^{W_u}, \\ y = Cx + a_y, \end{cases} \tag{13}$$

where the attack input $a_u^{W_u}$ can be regarded as an unknown input and the columns of b_{W_u} are b_i, $i \in W_u$. Denote by b_{J_u} the matrix whose columns are b_i for $i \in J_u$. Let q_1 and q_2 be the largest integers such that for all $J_u \subset \{1, \ldots, n_u\}$ with $\mathrm{card}(J_u) \leq 2q_1 < n_u$ and $J_s \subset \{1, \ldots, n_y\}$ with $\mathrm{card}(J_s) \geq n_y - 2q_2 > 0$, the following is satisfied:

($\mathbf{c_3}$) $\mathrm{rank}(C^{J_s} b_{J_u}) = \mathrm{rank}(b_{J_u}) = \mathrm{card}(J_u)$.

($\mathbf{c_4}$) There exists $(N_{J_{us}}, L_{J_{us}}, E_{J_{us}}, T_{J_{us}})$ satisfying:

$$\begin{cases} N_{J_{us}}(I - E_{J_{us}} C^{J_s}) + L_{J_{us}} C^{J_s} + (E_{J_{us}} C^{J_s} - I) A = \mathbf{0}, \\ (T_{J_{us}} + E_{J_{us}} C^{J_s} - I) B = \mathbf{0}, \\ (E_{J_{us}} C^{J_s} - I) b_{J_u} = \mathbf{0}, \end{cases} \tag{14}$$

with detectable pair $(C^{J_s}, A - E_{J_{us}} C^{J_s} A)$ and stable $N_{J_{us}}$. If conditions (c_3) and (c_4) are satisfied, a UIO with the following structure exists for each b_{J_u} with $J_u \subset \{1, \ldots, n_u\}$, $\mathrm{card}(J_u) \leq 2q_1 < n_u$ and each C^{J_s} with $J_s \subset \{1, \ldots, n_y\}$, $\mathrm{card}(J_s) \geq n_y - 2q_2 > 0$:

$$\begin{cases} z_{J_{us}}^+ = N_{J_{us}} z_{J_{us}} + T_{J_{us}} Bu + L_{J_{us}} y^{J_s}, \\ \hat{x}_{J_{us}} = z_{J_{us}} + E_{J_{us}} y^{J_s}, \end{cases} \tag{15}$$

where $z_{J_{us}} \in \mathbb{R}^n$ is the observer state, $z_{J_{us}}^+ \in \mathbb{R}^n$ is the observer state at the next time step, $\hat{x}_{J_{us}}$ denotes the state estimate, and $(N_{J_{us}}, L_{J_{us}}, T_{J_{us}}, E_{J_{us}})$ are observer matrices satisfying (14), see [31] for further details. That is, system (15) is a UIO for the system:

$$\begin{cases} x^+ = Ax + Bu + b_{J_u} a_u^{J_u}, \\ y^{J_s} = C^{J_s} x + a_y^{J_s}, \end{cases} \tag{16}$$

with unknown input $b_{J_u} a_u^{J_u}$ and known input Bu. It follows that the estimation error $e_{J_{us}} = \hat{x}_{J_{us}} - x$ satisfies:

$$e_{J_{us}}^+ = N_{J_{us}} e_{J_{us}}, \tag{17}$$

for some stable matrix $N_{J_{us}}$. We refer to UIOs of the form (16) as *partially unknown* UIOs for the pair (J_u, J_s).

Assumption 3 *There are at most q_1 actuators and at most q_2 sensors attacked, i.e.,*

$$\mathrm{card}(W_u) \leq q_1 < \frac{n_u}{2}, \tag{18}$$

$$\mathrm{card}(W_y) \leq q_2 < \frac{n_y}{2}, \tag{19}$$

where q_1 and q_2 are the largest positive integers satisfying (c_3) and (c_4).

Remark 6 *Note that if conditions (c_3) and (c_4) are satisfied for b_{J_u} with $\text{card}(J_u) = 2q_1 = n_u$, then conditions (c_1) and (c_2) are satisfied, and (15) is a completely UIO for (1) with $T_{J_{us}} = \mathbf{0}$. Since we are considering partially unknown UIOs, we assume $2q_1 < n_u$ to exclude this case.*

Lemma 2 *Under Assumption 3, for each set of q_1 actuators, among all its supersets with $2q_1$ actuators, at least one set is a superset of W_u.*

Lemma 3 *Under Assumption 3, among each set of $n_y - q_2$ sensors, at least $n_y - 2q_2 > 0$ sensors are attack-free.*

Proof: Lemmas 2 and 3 follow trivially from Assumption 3. ■

Note that the existence of a UIO for each pair (J_u, J_s) with $\text{card}(J_u) \leq 2q_1$ and $\text{card}(J_s) \geq n_y - 2q_2$ means that if $W_u \subseteq J_u$ and $a_y^{J_s}(k) = 0$ for all $k \geq 0$, the estimation error $e_{J_{us}} = \hat{x}_{J_{us}} - x$ satisfies

$$|e_{J_{us}}(k)| \leq c_{J_{us}} \lambda_{J_{us}}^k |e_{J_{us}}(0)|, \tag{20}$$

for some $c_{J_{us}} > 0$ and $\lambda_{J_{us}} \in (0,1)$, and all $e_{J_{us}}(0) \in \mathbb{R}^n$, $k \geq 0$. Let Assumption 3 be satisfied. We construct a UIO for each pair (J_u, J_s) with $\text{card}(J_u) = q_1$ and $\text{card}(J_s) = n_y - q_2$. Then, we construct a UIO for each pair (S_u, S_s) with $\text{card}(S_u) = 2q_1$, $S_u \subset \{1, \ldots, n_u\}$, and $\text{card}(S_s) = n_y - 2q_2$, $S_s \subset \{1, \ldots, n_y\}$. Under Assumption 3, there exists at least one set $\bar{J}_u$ with $\text{card}(\bar{J}_u) = q_1$ such that $W_u \subseteq \bar{J}_u$ and at least one set $\bar{J}_s$ with $\text{card}(\bar{J}_s) = n_y - q_2$ such that $a_y^{\bar{J}_s}(k) = 0$ for all $k \geq 0$. Then, the estimate given by the UIO for $(\bar{J}_u, \bar{J}_s)$ is a correct estimate, and the estimates given by the UIOs for any (S_u, S_s) (denoted as $\hat{x}_{S_{us}}$) with $\text{card}(S_u) = 2q_1$, $S_u \supset \bar{J}_u$, and $\text{card}(S_s) = n_y - 2q_2$, $S_s \subset \bar{J}_s$, are consistent with $\hat{x}_{J_{us}}$. This motivates the following estimation strategy.

For each pair (J_u, J_s) with $\text{card}(J_u) = q_1$ and $\text{card}(J_s) = n_y - q_2$, define $\pi_{J_{us}}(k)$ as the largest deviation between $\hat{x}_{J_{us}}(k)$ and $\hat{x}_{S_{us}}(k)$ that is given by any pair (S_u, S_s) satisfying $\text{card}(S_u) = 2q_1$, $S_u \supset J_u$, $\text{card}(S_s) = n_y - 2q_2$, and $S_s \subset J_s$. That is,

$$\pi_{J_{us}}(k) := \max_{S_u \supset J_u, S_s \subset J_s} |\hat{x}_{J_{us}}(k) - \hat{x}_{S_{us}}(k)|, \tag{21}$$

for all $k \geq 0$. Define the sequences $\sigma_u(k)$ and $\sigma_s(k)$ as

$$(\sigma_u(k), \sigma_s(k)) := \arg\min_{J_u, J_s} \ \pi_{J_{us}}(k). \tag{22}$$

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

$$\hat{x}(k) = \hat{x}_{\sigma_{us}(k)}(k), \tag{23}$$

is an exponential attack-free estimate of the system state. For simplicity and without loss of generality, for all (J_u, J_s, S_u, S_s), $z_{J_{us}}(0)$ and $z_{S_{us}}(0)$ are chosen such that $\hat{x}_{J_{us}}(0) = \hat{x}_{S_{us}}(0) = \hat{x}(0)$. The following result summarizes the ideas presented above.

Theorem 2 *Consider system (1), observer (15), and the partially unknown multi-observer estimator (21)-(23). Define the estimation error $e(k) := \hat{x}_{\sigma_{us}(k)}(k) - x(k)$ and let (c_3)-(c_4) and Assumptions 1, 3 be satisfied; then, there exist positive constants $\bar{c} > 0$ and $\bar{\lambda} \in (0,1)$ satisfying:*

$$|e(k)| \leq \bar{c}\bar{\lambda}^k |e(0)|, \tag{24}$$

for all $e(0) \in \mathbb{R}^n$, $k \geq 0$.

Proof: See Appendix B. ■

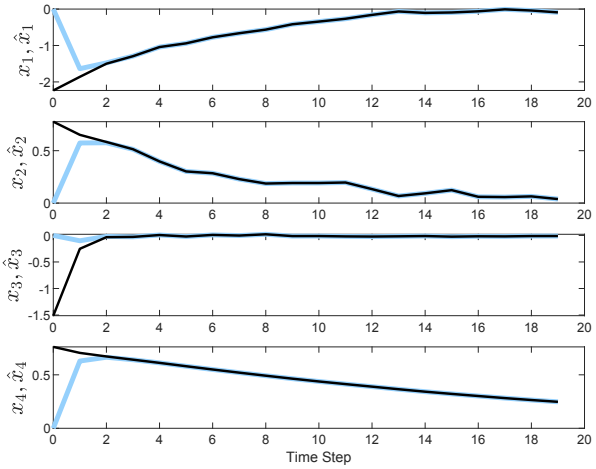

Fig. 3. Estimated states $\hat{x}$ converges to the true states x when $a_{u1} \sim \mathcal{U}(-1,1), a_{y4} \sim \mathcal{U}(-10,10)$. Legend: $\hat{x}$ (blue), true states (black)

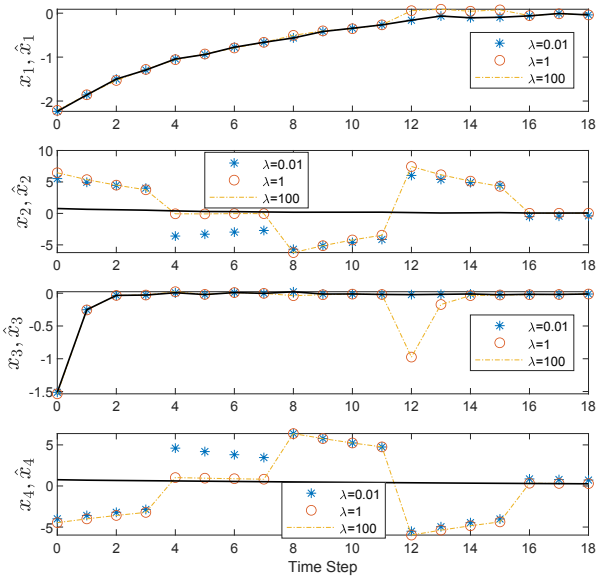

Fig. 4. Performance of estimator [20] for various values of λ, $a_{u1} \sim \mathcal{U}(-1,1), a_{y4} \sim \mathcal{U}(-10,10)$.

Example 2: The authors in [40], [41], [42] study the fault detection problem for a well stirred chemical reactor with a heat exchanger. This is a benchmark example used by the fault-detection community. We use this system to demonstrate our results. In particular, we consider the fourth-order model in [42, Section 6], which is of the form (1) with matrices (A, B)

given by

$$A = \begin{bmatrix} 0.8353 & 0 & 0 & 0 \\ 0 & 0.8324 & 0 & 0.0031 \\ 0 & 0.0001 & 0.1633 & 0 \\ 0 & 0.0280 & 0.0172 & 0.9320 \end{bmatrix}, \quad (25)$$

$$B = \begin{bmatrix} 0.0458 & 0 & 0 \\ 0 & 0.0457 & 0 \\ 0 & 0 & 0.0231 \\ 0 & 0.0007 & 0.0006 \end{bmatrix}.$$

The system [42, Section 6] uses full state measurements, i.e., $C = I_4$. However, for this C, it can be verified that neither completely nor partially UIOs exist for any submatrix C^{J_s}. So, for our scheme to work, we construct redundant virtual sensors using state measurements, and send virtual measurements through the unsecured network instead (see Remark 4 for details). We use the following C matrix:

$$y = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} x =: Cx.$$

It can be verified that completely UIOs do not exist for any C^{J_s} with $\mathrm{card}(J_s) \leq 2$. However, a partially UIO exists for each pair (J_u, J_s) with $\mathrm{card}(J_u) \leq 2$ and $\mathrm{card}(J_s) \geq 2$; then, $2q_1 = 2$ and $4 - 2q_2 = 2$, i.e., $q_1 = q_2 = 1$. We let $W_u = \{1\}$, $W_y = \{4\}$, i.e., the first actuator and the fourth sensor are attacked, $(u_1, u_2, u_3) \sim \mathcal{U}(-1, 1)$, and $a_{u1} \sim \mathcal{U}(-1, 1)$, $a_{y4} \sim \mathcal{U}(-10, 10)$, and $(x_1(0), x_2(0), x_3(0)) \sim \mathcal{N}(0, 1^2)$. We construct a partially UIO for each pair (J_u, J_s) with $\mathrm{card}(J_u) = 1, \mathrm{card}(J_s) = 3$ and each pair (S_u, S_s) with $\mathrm{card}(S_u) = 2, \mathrm{card}(S_s) = 2$. Therefore, totally $\binom{3}{1} \times \binom{4}{3} + \binom{3}{2} \times \binom{4}{2} = 30$ partially UIOs are designed. We initialize the observers at $\hat{x}(0) = [0, 0, 0]^\top$. Estimator (15), (21)-(23) is used to construct $\hat{x}(k)$. The performance of the estimator is shown in Figure 3. Similarly, for $k \in [0, 18]$, we use the estimator [20, Eq. (12)] for constructing $\hat{x}(k)$. Various values of the parameter λ in [20, Eq. (12)] are attempted. The performance of the estimator is shown in Figure 4. The execution time of the proposed multi-observer estimator to reconstruct $\hat{x}(k)$, $k \in [0, 19]$, is 0.262 seconds. For comparison, the estimator in [20, Eq. (12)] takes 7.960 seconds to compute the estimate for $k \in [0, 18]$.

IV. Attack Isolation and Reconstruction

Once we have an estimate $\hat{x}(k)$ of $x(k)$, either using the completely unknown multi-observer estimator in Section III-A or the partially unknown multi-observer estimator in Section III-B, we can use these estimates, the system model (1), and the known inputs to exponentially reconstruct the attack signals. Note that $e = \hat{x} - x \Rightarrow x = \hat{x} - e \Rightarrow x^+ = \hat{x}^+ - e^+$. Then, because B has full column rank (as introduced in the system description), the system dynamics (1) can be written in terms of e and $\hat{x}$ as follows:

$$\begin{cases} \hat{x}^+ = e^+ + A(\hat{x} - e) + B(u + a_u), \\ \Downarrow \\ a_u = B_{left}^{-1}(\hat{x}^+ - A\hat{x}) - u - B_{left}^{-1}(e^+ - Ae), \end{cases} \quad (26)$$

where B_{left}^{-1} denotes the Moore-Penrose pseudoinverse of B. Similarly, we have

$$\begin{cases} y = Cx + a_y = C\hat{x} - Ce + a_y, \\ \Downarrow \\ a_y = y - C\hat{x} + Ce. \end{cases} \quad (27)$$

First, consider the completely unknown multi-observer estimator in Section III-A. Let the estimation error dynamics (characterized by (6)-(8)) be given by

$$e^+ = f_1(e, x, a_y, a_u), \quad (28)$$

where $f_1 : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \to \mathbb{R}^n$ denotes some nonlinear function. That is, the estimation error is given by some nonlinear function of the state and the attack signals. However, in Theorem 1, we have proved that e converges to the origin exponentially. Hence, the terms depending on e and e^+ in the expression for a_u and a_y in (26) and (27) vanishes exponentially and therefore, the following attack estimates:

$$\hat{a}_u(k) := B_{left}^{-1}(\hat{x}(k) - A\hat{x}(k-1)) - u(k-1), \quad (29)$$

$$\hat{a}_y(k) := y(k) - C\hat{x}(k), \quad (30)$$

exponentially reconstruct the attack signals $a_u(k-1)$ and $a_y(k)$. That is,

$$\lim_{k\to\infty} (\hat{a}_u(k) - a_u(k-1)) = 0, \quad (31)$$

and

$$\lim_{k\to\infty} (\hat{a}_y(k) - a_y(k)) = 0. \quad (32)$$

Then, for sufficiently large k, the sparsity pattern of $\hat{a}_u(k)$ and $\hat{a}_y(k)$ can be used to isolate attacks, i.e.,

$$\hat{W}_u(k) := \mathrm{supp}(\hat{a}_u(k)), \quad (33)$$

and

$$\hat{W}_y(k) := \mathrm{supp}(\hat{a}_y(k)), \quad (34)$$

where $\hat{W}_u(k)$ denotes the set of isolated attacked actuators, and $\hat{W}_y(k)$ denotes the set of isolated attacked sensors. Note that we can only estimate a_u from $\hat{x}^+$ and e^+, which implies that we always have, at least, a one-step delay when isolating actuator attacks.

Next, consider the partially unknown multi-observer estimator given in Section III-B. In this case, the attack vector a_u and a_y can also be written as (26) and (27), and the estimation error dynamics is given by some nonlinear difference equation characterized by the estimator structure in (21)-(23). Let the estimation error dynamics be given by

$$e^+ = f_2(e, x, a_y, a_u), \quad (35)$$

for some nonlinear function $f_2 : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \to \mathbb{R}^n$. In Theorem 2, we have proved that e converges to the origin exponentially. Hence, the attack estimate in (29) and (30) exponentially reconstructs the attack signals. Again, the sparsity pattern of $\hat{a}_u(k)$ and $\hat{a}_y(k)$ can be used to isolate actuator and sensor attacks using (33) and (34).

Remark 7 *To reconstruct attack signals, we have to find the solution to a set of linear equations. However, this set of equations does not always have a unique solution. If the solution is not unique, there is an infinite number of attack*

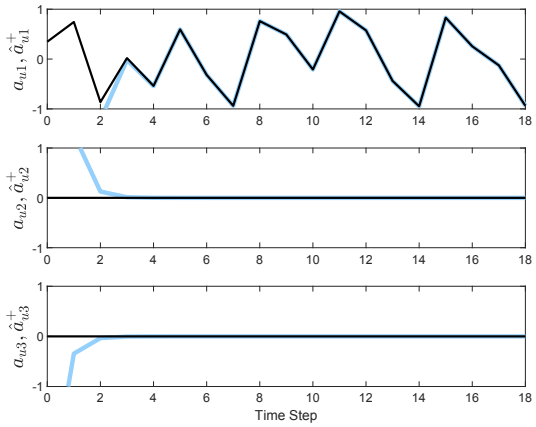

Fig. 5. Estimated actuator attacks $\hat{a}_u^+$ converges to a_u when $a_{u1} \sim \mathcal{U}(-1,1), a_{y4} \sim \mathcal{U}(-10,10)$. Legend: $\hat{a}_u^+$ (blue), a_u (black).

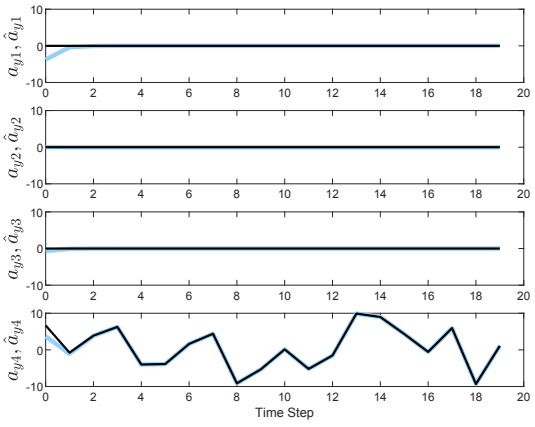

Fig. 6. Estimated sensor attacks $\hat{a}_y$ converges to a_y when $a_{u1} \sim \mathcal{U}(-1,1), a_{y4} \sim \mathcal{U}(-10,10)$. $\hat{a}_y$ (blue), a_y (black).

vectors a_u that explain the data, and thus it is impossible to find the correct attack vector. For instance, if we have a linear system, $x^+ = Ax + Ba_u$, and we know (x^+, x, A, B), there is a unique solution, $a_u = a_u^*$, to the set of linear equations, $x^+ - Ax = Ba_u$, if and only if the matrix B has full column rank. When B does not have full column rank, B has a nontrivial kernel, and thus there is an infinite number of solutions, $a_u = a_u^*$, to $x^+ - Ax = Ba_u$. The latter means that we cannot reconstruct the attack vector from (x^+, x, A, B). This is the fundamental reason why we need to assume that B has full column rank (see (26)*). Assuming full column rank B is a standard assumption in the fault diagnosis and security of control systems literature, see, e.g.,* [43], [44], [45].

Remark 8 *The attack isolation algorithm we provide is sensitive to process disturbance and measurement noise since the estimate of the attack signals we obtain from equations* (29)-(30) *will be an estimate of the mixture of attack signals and noise, which is unlikely to be sparse anymore. Isolation algorithms design for noisy systems is an interesting problem for future work.*

Example 4: Here we consider system (25) and the partially unknown multi-observer estimator in Example 2. Let $W_u = \{1\}$, $W_y = \{4\}$, $(u_1, u_2, u_3) \sim \mathcal{U}(-1,1)$, $a_{u1} \sim \mathcal{U}(-1,1)$, $a_{y4} \sim \mathcal{U}(-10,10)$, and $(x_1(0), x_2(0), x_3(0)) \sim \mathcal{N}(0, 1^2)$. We obtain $\hat{a}_u(k)$ and $\hat{a}_y(k)$ from (29) and (30). The reconstructed attacks are shown in Figures 5-6. In this case, using sparsity of the estimated attacks, actuator 1 and sensor 4 are correctly isolated.

V. Control

In this section, we introduce a method to use the proposed multi-observer estimators to asymptotically stabilize the system dynamics.

A. Sensor attacks only

We first consider the case when only sensors are attacked and actuators are attack-free. Then, the system is given by

$$\begin{cases} x^+ = Ax + Bu, \\ y = Cx + a_y. \end{cases} \tag{36}$$

Let $u = K\hat{x}$, where $\hat{x}$ is the estimate given by the estimator in Section III-A or the estimator in Section III-B, and K is chosen such that $A + BK$ is stable. Then, the closed-loop system is given by

$$x^+ = Ax + BK\hat{x}, \tag{37}$$

or in terms of the estimation error as

$$\begin{aligned} x^+ &= Ax + B(K(\hat{x} - x + x)), \\ &= (A + BK)x + BKe. \end{aligned} \tag{38}$$

For the completely unknown multi-observer estimator, let the estimation error dynamics be given by

$$e^+ = f_1(e, x, a_y), \tag{39}$$

for some nonlinear function $f_1 : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n_y} \to \mathbb{R}^n$. For the partially unknown multi-observer estimator, let the estimation error dynamics be given by

$$e^+ = f_2(e, x, a_y), \tag{40}$$

for some nonlinear function $f_2 : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n_y} \to \mathbb{R}^n$. Since $A + BK$ is stable, the closed-loop dynamics (38) is Input-to-State Stable (ISS) with respect to input $e(k)$ and some linear gain, see [46] for details. Moreover, in Theorems 1 and 2, we have proved that (39) and (40) are exponentially stable uniformly in $x(k)$ and $a_y(k)$. The latter and ISS of the system dynamics imply that $\lim_{k\to\infty} x(k) = 0$ [46].

B. Sensor and actuator attacks

Here, we consider sensor and actuator attacks. We propose a simple yet effective technique to stabilize the system by switching off the isolated attacked actuators, i.e., by removing

the columns of B that correspond to the isolated actuators, and closing the loop with a multi-observer based output dynamic feedback controller. We introduce a switching signal $\rho(k) \subseteq \{1,\dots n_u\}$, containing the isolated attack-free actuators, i.e., $\rho(k) := \{1,\dots,n_u\} \setminus \hat{W}_u(k)$. This $\rho(k)$ is used to denote actuators that are kept switched on at time k. That is, $\rho(k) = J$ if the subset $J \subseteq \{1,\dots,n_u\}$ of actuators are switched on and the remaining actuators are switched off at time k. Here, we are assuming that there is a secure communication channel that can be used to transmit $\rho(k)$ to the actuators of the system so that, once $\rho(k)$ has arrived, we could physically turn the isolated actuators off. Again, let B be partitioned as $B = [b_1,\dots,b_i,\dots,b_{n_u}]$. After switching off the subset $\{1,\dots,n_u\}\setminus\rho(k)$ of actuators, system (1) is written as follows

$$\begin{cases} x^+ = Ax + b_{\rho(k)}(u^{\rho(k)} + a_u^{\rho(k)}), \\ y = Cx + a_y, \end{cases} \tag{41}$$

where $b_{\rho(k)}$ is the matrix whose columns are $b_i, i \in \rho(k)$, vectors $u^{\rho(k)}$ and $a_u^{\rho(k)}$ are the inputs and attacks corresponding to the switched-on actuators, respectively. We first consider the case when the completely unknown multi-observer estimator in Section III-A exists, i.e., $\hat{x}$ is generated by (6)-(8). We estimate $\hat{a}_u(k)$ using (29) and obtain $\hat{W}_u(k)$ from (33). Then, we switch off the set $\hat{W}_u$ of actuators by letting $\rho(k) = \bar{J}(k) = \{1,\dots,n_u\} \setminus \hat{W}_u(k)$. Since $a_{ui}(k) = 0, i \in \bar{J}(k)$, system (41) has the following form:

$$x^+ = Ax + b_{\bar{J}(k)} u^{\bar{J}(k)} \tag{42}$$

where $u^{\bar{J}(k)} \in \mathbb{R}^{\mathrm{card}(\bar{J}(k))}$ is the set of isolated attack-free inputs. Let $0 < q^\star < n_u$ be the largest integer such that (A, b_J) is stabilizable for each set $J \subset \{1,\dots,n_u\}$ with $\mathrm{card}(J) \geq n_u - q^\star$ where b_J denotes a matrix whose columns are b_i for $i \in J$. We assume that at most $q^\star$ actuators are attacked. It follows that $n_u - q^\star \leq \mathrm{card}(\bar{J}(k)) \leq n_u$. We assume the following.

Assumption 4 *For any subset J with cardinality $\mathrm{card}(J) = n_u - q^\star$, there exists a linear switching state feedback controller $u^{\bar{J}(k)} = K_{\bar{J}(k)}x$ such that the closed-loop dynamics:*

$$x^+ = (A + b_{\bar{J}(k)}K_{\bar{J}(k)})x + b_{\bar{J}(k)}K_{\bar{J}(k)}e, \tag{43}$$

is ISS with respect to input e for $b_{\bar{J}(k)}$ arbitrarily switching among all $b_{J'}$ with $J \subset J' \subset \{1,\dots,n_u\}$ and $n_u - q^\star \leq \mathrm{card}(J') \leq n_u$.

Remark 9 *We do not give a method for designing the linear switching state feedback controller $u^{\bar{J}(k)} = K_{\bar{J}(k)}x$. Standard results for designing switching controllers, for instance results in* [47] *and references therein, can be used to design controllers satisfying Assumption 4.*

By switching off the set $\hat{W}_u(k)$ of actuators at time k, using the controller designed for the set $\bar{J}(k)$, and letting $u^{\bar{J}(k)} = K_{\bar{J}(k)}\hat{x}$, the closed-loop system can be written as (43) with estimation error $e = \hat{x} - x$ generated by some nonlinear difference equation (28). Because in Theorem 1, we have proved that $e(k)$ converges to zero exponentially uniformly in $x(k)$, $a_y(k)$ and $a_u(k)$, the error $e(k)$ in (43) is a vanishing perturbation. Hence, under Assumption 4, it follows that $\lim_{k\to\infty} x(k) = 0$.

Next, assume that a completely unknown multi-observer estimator does not exist but a partially unknown multi-observer estimator exists (Section III-B), i.e., $\hat{x}$ is generated from (21)-(23) and $q_1 \leq q^\star$. We assume that at most q_1 actuators are attacked. We construct $\hat{x}(k)$ from (21)-(23), estimate $\hat{a}_u(k)$ using (29), and obtain $\hat{W}_u(k)$ from (33). After switching off the set $\hat{W}_u(k)$ of actuators, the system has the form (42) with $n_u - q_1 \leq \mathrm{card}(\bar{J}(k)) \leq n_u$. We assume the following.

Assumption 5 *For any subset J with cardinality $\mathrm{card}(J) = n_u - q_1$, there exists a linear switching state feedback controller $u^{\bar{J}(k)} = K_{\bar{J}(k)}x$ such that the closed-loop dynamics* (43) *is ISS with respect to input e for $b_{\bar{J}(k)}$ arbitrarily switching among all $b_{J'}$ with $J \subset J' \subset \{1,\dots,n_u\}$ and $n_u - q_1 \leq \mathrm{card}(J') \leq n_u$.*

Using the controller designed for the set $\bar{J}(k)$, and letting $u^{\bar{J}(k)} = K_{\bar{J}(k)}\hat{x}$, the closed-loop dynamics can be written in the form (43). Then, in this case, $e(k)$ is generated by some nonlinear difference equation of the form (35). Under Assumption 5, the closed-loop dynamics (43) is ISS with input $e(k)$, see [46]. Moreover, in Theorem 2, we have proved that $e(k)$ converges to the origin exponentially uniformly in $x(k)$, $a_u(k)$ and $a_y(k)$. The latter and ISS of the system dynamics imply that $\lim_{k\to\infty} x(k) = 0$ [46].

Example 6: Consider the following system:

$$\begin{cases} x^+ = \begin{bmatrix} 0.5 & 0 & 0.1 \\ 0.2 & 1.7 & 0 \\ 1 & 0 & 0.3 \end{bmatrix} x + \begin{bmatrix} 0.5 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} (u + a_u), \\ y = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix} x + a_y. \end{cases} \tag{44}$$

Since (A, b_i) is stabilizable for $i \in \{1,2,3\}$, we have $q^\star = 2$. It can be verified that there does not exist a completely UIO for any $S_s \subset \{1,2,3,4\}$ with $\mathrm{card}(S_s) = 2$, but partially UIOs exists for each pair (J_u, J_s) with $\mathrm{card}(J_u) \leq 2$ and $\mathrm{card}(J_s) \geq 2$; then, we have $q_1 = q_2 = 1$ and $q_1 < q^\star$. We let $W_u = \{3\}$, $W_y = \{2\}$, and $a_{u3}, a_{y2} \sim \mathcal{U}(-10, 10)$. We construct $\binom{3}{1} \times \binom{4}{3} + \binom{3}{2} \times \binom{4}{2} = 30$ UIOs and use the design method given in [47] to build controllers for actuators $\{1,2\}$, $\{1,3\}$, $\{2,3\}$, $\{1,2,3\}$. Then, we use the partially unknown multi-observer approach in Section III-B to estimate the state, reconstruct the attack signals and control the system. The state of the system is shown in Figure 7.

VI. Conclusion

We have addressed the problem of state estimation, attack isolation, and control for discrete-time linear time-invariant (LTI) systems under (potentially unbounded) actuator and sensor false data injection attacks. Using a bank of Unknown Input Observers (UIOs), we have proposed an estimator that reconstructs the system state and the attack signals. These estimates are then used to isolate attacks and stabilize the system dynamics. We have proposed an effective technique to stabilize the system by switching off isolated actuators.

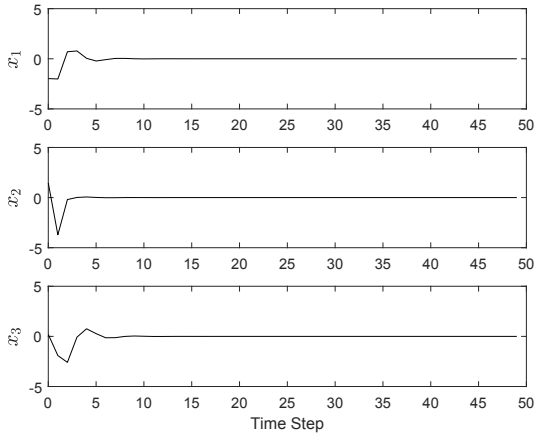

Fig. 7. State trjectories when $a_{u3}, a_{y2} \sim \mathcal{U}(-10, 10)$.

Possible future directions include considering a way to reduce the number of observers required for the estimation scheme, extending the scheme for systems that are subject to noise and system disturbances, and exploring the idea of redundant communication channels and virtual sensors to reduce the computational complexity of the scheme.

APPENDIX A
PROOF OF THEOREM 1

Under Assumption 2, there exists at least one set $\bar{J}_s$ with $\text{card}(\bar{J}_s) = n_y - q$ such that $a_y^{\bar{J}_s}(k) = 0$ for all $k \geq 0$. Then, there exist $c_{\bar{J}_s} > 0$ and $\lambda_{\bar{J}_s} \in (0, 1)$ such that

$$|e_{\bar{J}_s}(k)| \leq c_{\bar{J}_s} \lambda_{\bar{J}_s}^k |e(0)|, \tag{45}$$

for all $e(0) \in \mathbb{R}^n$ and $k \geq 0$. Moreover, for any set $S_s \subset \bar{J}_s$ with $\text{card}(S_s) = n_y - 2q$, we have $a_y^{S_s}(k) = 0 \; \forall k \geq 0$; hence, there exist $c_{S_s} > 0$ and $\lambda_{S_s} \in (0, 1)$ such that

$$|e_{S_s}(k)| \leq c_{S_s} \lambda_{S_s}^k |e(0)|, \tag{46}$$

for all $e(0) \in \mathbb{R}^n$ and $k \geq 0$. Consider π_{J_s} in (6). Combining the above inequalities, we have

$$\begin{aligned} \pi_{\bar{J}_s}(k) =& \max_{S_s \subset \bar{J}_s} |\hat{x}_{\bar{J}_s}(k) - \hat{x}_{S_s}(k)| \\ =& \max_{S_s \subset \bar{J}_s} |\hat{x}_{\bar{J}_s}(k) - x(k) + x(k) - \hat{x}_{S_s}(k)| \\ \leq& |e_{\bar{J}_s}(k)| + \max_{S_s \subset \bar{J}_s} |e_{S_s}(k)| \\ \leq& 2c'_{\bar{J}_s} \lambda'^k_{\bar{J}_s} |e(0)|, \end{aligned} \tag{47}$$

for all $e(0) \in \mathbb{R}^n$ and $k \geq 0$, where

$$c'_{\bar{J}_s} := \max_{S_s \subset \bar{J}_s} \left\{ c_{\bar{J}_s}, c_{S_s} \right\},$$

$$\lambda'_{\bar{J}_s} := \max_{S_s \subset \bar{J}_s} \left\{ \lambda_{\bar{J}_s}, \lambda_{S_s} \right\}.$$

Note that $S_s \subset \bar{J}_s$, $\text{card}(S_s) = n_y - 2q$. Then, from (7), we have $\pi_{\sigma_s(k)}(k) \leq \pi_{\bar{J}_s}(k)$. From Lemma 1, we know that there exist at least one set $\bar{S}_s \subset \sigma_s(k)$ with $\text{card}(\bar{S}_s) = n_y - 2q$, such that $a_y^{\bar{S}_s}(k) = 0$ for all $k \geq 0$, and there exist $c_{\bar{S}_s} > 0$ and $\lambda_{\bar{S}_s} \in (0, 1)$ such that

$$|e_{\bar{S}_s}(k)| \leq c_{\bar{S}_s} \lambda_{\bar{S}_s}^k |e(0)|, \tag{48}$$

for all $e(0) \in \mathbb{R}^n$ and $k \geq 0$. From (6), we have

$$\begin{aligned} \pi_{\sigma_s(k)}(k) =& \max_{S_s \subset \sigma_s(k)} |\hat{x}_{\sigma_s(k)}(k) - \hat{x}_{S_s}(k)| \\ \geq& |\hat{x}_{\sigma_s(k)}(k) - \hat{x}_{\bar{S}_s}(k)|. \end{aligned}$$

Using this lower bound on $\pi_{\sigma_s(k)}(k)$ and the triangle inequality we have that

$$\begin{aligned} |e_{\sigma_s(k)}(k)| =& |\hat{x}_{\sigma_s(k)}(k) - x(k)| \\ =& |\hat{x}_{\sigma_s(k)}(k) - \hat{x}_{\bar{S}_s}(k) + \hat{x}_{\bar{S}_s}(k) - x(k)| \\ \leq& |\hat{x}_{\sigma_s(k)}(k) - \hat{x}_{\bar{S}_s}(k)| + |e_{\bar{S}_s}(k)| \\ \leq& \pi_{\sigma_s(k)}(k) + |e_{\bar{S}_s}(k)| \\ \leq& \pi_{\bar{J}_s}(k) + |e_{\bar{S}_s}(k)|, \end{aligned} \tag{49}$$

for all $k \geq 0$. Hence, from (47) and (48), we have

$$|e_{\sigma_s(k)}(k)| \leq \bar{c} \bar{\lambda}^k |e(0)|, \tag{50}$$

for all $e(0) \in \mathbb{R}^n$ and $k \geq 0$, where $\bar{c} = 3 \max\{c_{\bar{S}_s}, c'_{\bar{J}_s}\}$ and $\bar{\lambda} = \max\{\lambda_{\bar{S}_s}, \lambda'_{\bar{J}_s}\}$. Inequality (50) is of the form (9) and the result follows.

APPENDIX B
PROOF OF THEOREM 2

Under Assumption 3, there exists at least one set $\bar{J}_u$ with $\text{card}(\bar{J}_u) = q_1$ such that $\bar{J}_u \supset W_u$, and at least one set $\bar{J}_s$ with $\text{card}(\bar{J}_s) = n_y - q_2$ such that $a_y^{\bar{J}_s}(k) = 0$ for all $k \geq 0$; then, there exist $c_{\bar{J}_{us}} > 0$ and $\lambda_{\bar{J}_{us}} \in (0, 1)$ satisfying

$$|e_{\bar{J}_{us}}(k)| \leq c_{\bar{J}_{us}} \lambda_{\bar{J}_{us}}^k |e(0)|, \tag{51}$$

for all $e(0) \in \mathbb{R}^n$ and $k \geq 0$. Moreover, for any set $S_u \supset \bar{J}_u$ with $\text{card}(S_u) = 2q_1$ and $S_s \subset \bar{J}_s$ with $\text{card}(S_s) = n_y - 2q_2$, we have $S_u \supset W_u$ and $a_y^{S_s}(k) = 0$ for all $k \geq 0$; hence, there exist $c_{S_{us}} > 0$ and $\lambda_{S_{us}} \in (0, 1)$ such that

$$|e_{S_{us}}(k)| \leq c_{S_{us}} \lambda_{S_{us}}^k |e(0)|, \tag{52}$$

for all $e(0) \in \mathbb{R}^n$ and $k \geq 0$. Consider $\pi_{\bar{J}_{us}}$ in (21).

Combining the above results, we have that

$$\begin{aligned} \pi_{\bar{J}_{us}}(k) =& \max_{S_u \supset \bar{J}_u, S_s \subset \bar{J}_s} |\hat{x}_{\bar{J}_{us}}(k) - \hat{x}_{S_{us}}(k)| \\ =& \max_{S_u \supset \bar{J}_u, S_s \subset \bar{J}_s} |\hat{x}_{\bar{J}_{us}}(k) - x(k) + x(k) - \hat{x}_{S_{us}}(k)| \\ \leq& |e_{\bar{J}_{us}}(k)| + \max_{S_u \supset \bar{J}_u, S_s \subset \bar{J}_s} |e_{S_{us}}(k)|, \end{aligned}$$

for all $k \geq 0$. From (51) and (52), we obtain

$$\pi_{\bar{J}_{us}}(k) \leq 2c'_{\bar{J}_{us}} \lambda'^k_{\bar{J}_{us}} |e(0)|, \tag{53}$$

for all $e(0) \in \mathbb{R}^n$ and $k \geq 0$, where

$$c'_{\bar{J}_{us}} := \max_{S_u \supset \bar{J}_u, S_s \subset \bar{J}_s} \left\{ c_{\bar{J}_{us}}, c_{S_{us}} \right\},$$

$$\lambda'_{\bar{J}_{us}} := \max_{S_u \supset \bar{J}_u, S_s \subset \bar{J}_s} \left\{ \lambda_{\bar{J}_{us}}, \lambda_{S_{us}} \right\}.$$

Note that $S_u \supset \bar{J}_u$, $\text{card}(S_u) = 2q_1$, and $S_s \subset \bar{J}_s$, $\text{card}(S_s) = n_y - 2q_2$. Then, from (22), we have $\pi_{\sigma_{us}(k)}(k) \leq \pi_{\bar{J}_{us}}(k)$. By Lemmas 2 and 3, we know that there exists at least one set $\bar{S}_u \supset \sigma_u(k)$ with $\text{card}(\bar{S}_u) = 2q_1$ and at least one set $\bar{S}_s \subset \sigma_s(k)$ with $\text{card}(\bar{S}_s) = n_y - 2q_2$ such that

$\bar{S}_u \supset W_u$ and $a_y^{\bar{S}_s}(k) = 0$ for all $k \geq 0$. Hence, there exist $c_{\bar{S}_{us}} > 0$ and $\lambda_{\bar{S}_{us}} \in (0,1)$ satisfying

$$|e_{\bar{S}_{us}}(k)| \leq c_{\bar{S}_{us}} \lambda_{\bar{S}_{us}}^k |e(0)|, \tag{54}$$

for all $e(0) \in \mathbb{R}^n$ and $k \geq 0$. From (21), by construction

$$\begin{aligned}\pi_{\sigma_{us}(k)}(k) =& \max_{S_u \supset \sigma_u(k), S_s \subset \sigma_s(k)} |\hat{x}_{\sigma_{us}(k)}(k) - \hat{x}_{S_{us}}(k)| \\ \geq& |\hat{x}_{\sigma_{us}(k)}(k) - \hat{x}_{\bar{S}_{us}}(k)|.\end{aligned}$$

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

$$\begin{aligned}|e_{\sigma_{us}(k)}(k)| =& |\hat{x}_{\sigma_{us}(k)}(k) - x(k)| \\ =& |\hat{x}_{\sigma_{us}(k)}(k) - \hat{x}_{\bar{S}_{us}}(k) + \hat{x}_{\bar{S}_{us}}(k) - x(k)| \\ \leq& |\hat{x}_{\sigma_{us}(k)}(k) - \hat{x}_{\bar{S}_{us}}(k)| + |e_{\bar{S}_{us}}(k)| \\ \leq& \pi_{\sigma_{us}(k)}(k) + |e_{\bar{S}_{us}}(k)| \\ \leq& \pi_{\bar{J}_{us}}(k) + |e_{\bar{S}_{us}}(k)|,\end{aligned} \tag{55}$$

for all $k \geq 0$. Hence, from (53) and (54), we have

$$|e_{\sigma_{us}(k)}(k)| \leq \bar{c}\bar{\lambda}^k |e(0)|, \tag{56}$$

for all $e(0) \in \mathbb{R}^n$ and $k \geq 0$, where $\bar{c} = 3\max\{c_{\bar{S}_{us}}, c'_{\bar{J}_{us}}\}$, $\bar{\lambda} = \max\{\lambda_{\bar{S}_{us}}, \lambda'_{\bar{J}_{us}}\}$. Inequality (56) is of the form (24), and the result follows.

REFERENCES

[1] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Transactions on Information and System Security*, vol. 14, no. 1, pp. 21–32, 2009.

[2] 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.

[3] 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.

[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)*, pp. 5906–5911, 2016.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

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

[10] 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.

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

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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*, 2018.

[17] T. Yang, C. Murguia, M. Kuijper, and D. Nešić, "Attack detection and isolation for discrete-time nonlinear systems," *2018 Australian & New Zealand Control Conference (ANZCC)*, 2018.

[18] S. M. Djouadi, A. M. Melin, E. M. Ferragut, J. A. Laska, J. Dong, and A. Drira, "Finite energy and bounded actuator attacks on cyber-physical systems," *2015 European Control Conference, ECC 2015*, pp. 3659–3664, 2015.

[19] X. Jin, W. M. Haddad, and T. Yucelen, "An adaptive control architecture for mitigating sensor and actuator attacks in cyber-physical systems," *IEEE Transactions on Automatic Control*, vol. 62, pp. 6058–6064, Nov 2017.

[20] 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.

[21] M. Showkatbakhsh, Y. Shoukry, R. H. Chen, S. Diggavi, and P. Tabuada, "An SMT-based approach to secure state estimation under sensor and actuator attacks," *2017 IEEE 56th Annual Conference on Decision and Control, CDC 2017*, 2017.

[22] Y. Shoukry, P. Nuzzo, A. Puggelli, A. L. 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, pp. 4917–4932, 2017.

[23] Y. Mo and B. Sinopoli, "Secure Estimation in the Presence of Integrity Attacks," *IEEE Transactions on Automatic Control*, vol. 60, no. 4, pp. 1145–1151, 2015.

[24] J. M. Moya, Á. Araujo, Z. Banković, J. M. de Goyeneche, J. C. Vallejo, P. Malagón, D. Villanueva, D. Fraga, E. Romero, and J. Blesa, "Improving security for SCADA sensor networks with reputation systems and self-organizing maps," *Sensors*, vol. 9, no. 11, pp. 9380–9397, 2009.

[25] M. Darms, P. Rybski, and C. Urmson, "Classification and tracking of dynamic objects with multiple sensors for autonomous driving in Urban environments," *IEEE Intelligent Vehicles Symposium, Proceedings*, pp. 1197–1202, 2008.

[26] F. V. Wyk, Y. Wang, A. Khojandi, and N. Masoud, "Real-Time Sensor Anomaly Detection and Identification in Automated Vehicles," pp. 1–13, 2019.

[27] Q. Liu, Y. Mo, X. Mo, C. Lv, E. Mihankhah, and D. Wang, "Secure Pose Estimation for Autonomous Vehicles under Cyber Attacks," *Intelligent Vehicles Symposium*, no. Iv, pp. 1401–1406, 2019.

[28] A. Laszka, W. Abbas, Y. Vorobeychik, and X. Koutsoukos, "Synergic security for smart water networks: Redundancy, diversity, and hardening," *Proceedings - 2017 3rd International Workshop on Cyber-Physical Systems for Smart Water Networks, CySWATER 2017*, pp. 21–24, 2017.

[29] A. Preis and A. Ostfeld, "Multiobjective contaminant response modeling for water distribution systems security," *Journal of Hydroinformatics*, vol. 10, no. 4, pp. 267–274, 2008.

[30] E. D. Sontag, "Input to state stability: Basic concepts and results," *Lecture Notes in Mathematics*, vol. 1932, pp. 163–220, 2008.

[31] S. X. Ding, *Model-based Fault Diagnosis Techniques: Design Schemes, Algorithms, and Tools*. Springer, 2013.

[32] M. Showkatbakhsh, Y. Shoukry, R. H. Chen, S. Diggavi, and P. Tabuada, "An SMT-based approach to secure state estimation under sensor and actuator attacks," *2017 IEEE 56th Annual Conference on Decision and Control, CDC 2017*, pp. 157–162.

[33] Y. Shoukry and P. Tabuada, "Event-triggered state observers for sparse sensor noise/attacks," *IEEE Transactions on Automatic Control*, vol. 61, no. 8, pp. 2079–2091, 2016.

[34] Y. Shoukry, M. Chong, M. Wakaiki, P. Nuzzo, A. L. Sangiovanni-Vincentelli, S. A. Seshia, J. P. Hespanha, and P. Tabuada, "SMT-Based Observer Design for Cyber-Physical Systems under Sensor Attacks,"

2016 ACM/IEEE 7th International Conference on Cyber-Physical Systems, ICCPS 2016 - Proceedings, 2016.
[35] C. Wu, Z. Hu, J. Liu, and L. Wu, "Secure estimation for cyber-physical systems via sliding mode," *IEEE Transactions on Cybernetics*, vol. 48, no. 12, pp. 3420–3431, 2018.
[36] E. Chong and S. Zak, *An Introduction to Optimization*. Wiley Series in Discrete Mathematics and Optimization, Wiley, 2013.
[37] J. Ploeg, N. van de Wouw, and H. Nijmeijer, "Lp string stability of cascaded systems : application to vehicle platooning," *IEEE Transactions on Control Systems Technology, accepted*.
[38] S. nc, J. Ploeg, N. van de Wouw, and H. Nijmeijer, "Cooperative adaptive cruise control: Network-aware analysis of string stability," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, pp. 1527–1537, 2014.
[39] Y. A. Harfouch, S. Yuan, and S. Baldi, "An adaptive switched control approach to heterogeneous platooning with intervehicle communication losses," *IEEE Transactions on Control of Network Systems*, vol. 5, pp. 1434–1444, 2018.
[40] J. Chen and R. Patton, *Robust Model-Based Fault Diagnosis for Dynamic Systems*. Springer Publishing Company, Incorporated, 2012.
[41] K. Watanabe and D. M. Himmelblau, "Fault diagnosis in nonlinear chemical processes. part ii. application to a chemical reactor," *AIChE Journal*, vol. 29, 1983.
[42] C. Murguia and J. Ruths, "On model-based detectors for linear time-invariant stochastic systems under sensor attacks," *IET Control Theory Applications*, vol. 13, pp. 1051–1061, 2019.
[43] W. Chen and M. Saif, "An Actuator Fault Isolation Strategy for Linear and Nonlinear Systems," *2005 American Control ConferenceJune 8-10, 2005. Portland, OR, USA*, pp. 3321–3326, 2005.
[44] W. Chen and M. Saif, "Actuator fault diagnosis for uncertain linear systems using a high-order sliding-mode robust differentiator (HOSMRD)," *INTERNATIONAL JOURNAL OF ROBUST AND NONLINEAR CONTROL*, vol. 18, no. April 2007, pp. 413–426, 2008.
[45] I. Shames, M. H. Teixeira, H. Sandberg, and K. H. Johansson, "Distributed Fault Detection for Interconnected Second-Order Systems," no. April, 2011.
[46] Z.-P. Jiang and Y. Wang, "Input-to-state stability for discrete-time nonlinear systems," *Automatica*, vol. 37, no. 6, pp. 857–869, 2001.
[47] J. Daafouz, P. Riedinger, and C. Iung, "Stability Analysis and Control Synthesis for Switched Systems: A switched Lyapunov function approach," *IEEE Trans. on Automat. Contr.*, vol. 47, no. 11, pp. 1883–1887, 2002.

Tianci Yang received the bachelor's degree of Engineering from the Honors school, Harbin Institute of Technology, Harbin, Heilongjiang, China, in 2014. He received the Ph.D. degree from the department of Electrical and Electronic Engineering, University of Melbourne, Melbourne, VIC, Australia, in 2019. He is currently a Research Fellow with the School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore. His current research interests include systems and control, security of cyber-physical systems.

Carlos Murguia was born in Mexico City, Mexico, in 1984. He received a Master of Science (MSc) degree (2015) in Electrical Engineering from the Center for Research and Advanced Studies of the National Polytechnic Institute (CINVESTAV), Mexico City, Mexico, in 2010. From November 2010 to May 2015, he was affiliated with the Dynamics and Control group at Eindhoven University of Technology, The Netherlands, where he pursued a Ph.D. degree in Mechanical Engineering. He has served as postdoctoral research fellow (2015–2019) at the Singapore University of Technology and Design (Singapore); the University of Melbourne (Australia); and the University of California (USA), Los Angeles. As of March 2020, he is an Assistant Professor in Mechanical Engineering Department at Eindhoven University of Technology. His research interests include nonlinear dynamics and control, mechatronics, security and privacy of cyberphysical systems, distributed control systems, networked control systems, fault detection, control of networks, stochastic systems, and time-delayed systems.

Margreta Kuijper is a Professor at the Department of Electrical and Electronic Engineering of the University of Melbourne (Australia) where she has been employed since 1995. From 1992 to 1995 she was a postdoctoral fellow at the Mathematics Department of the University of Groningen, the Netherlands. From 1988 to 1992 she worked at the Centrum Wiskunde & Informatica (CWI), Amsterdam, where she obtained her PhD degree in 1992. Her main research interests evolve around the interplay between coding theory and systems theory. Current research interests include algebraic error control coding and cybersecurity of control systems.

Dragan Nešić (F'08) received the Ph.D. degree in systems engineering from the Research School of Information Sciences, Australian National University, Canberra, ACT, Australia, in 1997. Since 1999, he has been with The University of Melbourne, Melbourne, VIC, Australia, where he is currently a Professor with the Department of Electrical and Electronic Engineering.

Dr. Nešić served as a member for the Board of Governors, CSS, and as a General Co-Chair of the IEEE CDC 2017. He is a recipient of Humboldt Research Award in 2020, and was a co-recipient of the George S. Axelby Outstanding Paper Award in 2018. He is a recipient of IFAC fellowship in 2020, and was a recipient of Humboldt Research Fellowship in 2003 by the Alexander von Humboldt Foundation, an Australian Professorial Fellowship from 2004 to 2009, and Future Fellowship from 2010 to 2014 by the Australian Research Council. He also served as an Associate Editor for the journals Automatica, the IEEE Transactions on Automatic Control, Systems and Control Letters, European Journal of Control.