Resilient Reinforcement in Secure State Estimation against Sensor Attacks with *a priori* Information

Takumi Shinohara, Toru Namerikawa, *Senior Member, IEEE*, and Zhihua Qu, *Fellow, IEEE*

***Abstract*—Recent control systems are severely depending on information technology infrastructures, especially the Internet of Things (IoT) devices, which create many opportunities for the interaction between physical world and cyberspace. Due to the tight connection, however, the cyber attacks have potentials to generate evil consequences to the physical entities, and therefore, securing the control systems is a vital issue for building smart societies. To this end, this paper especially deals with the state estimation problem in the presence of malicious sensor attacks. Unlike the existing work, in this paper, we consider the problem with *a priori* information of the state to be estimated. Specifically, we address three prior knowledge: the sparsity information, $(\alpha, \bar{n}_0)$-sparsity information, and side information, and in each scenario, we show that the state can be reconstructed even if more sensors are compromised. This implies that the prior information reinforces the system resilience against malicious sensor attacks. Then, an estimator under sensor attacks considering the information is developed and, under a certain condition, the estimator can be relaxed into a tractable convex optimization problem. Further, we extend this analysis to systems in presence of measurement noises, and it is shown that the prior information reduces the state-estimation error caused by the noise. The numerical simulations in a diffusion process finally illustrate the reinforcement and error-reduction results with the information.**

***Index Terms*—System security, secure state estimation, sensor attacks.**

I. Introduction

A. Background and Related Work

ADVANCES in networks and information technologies have great potential for the building of smart societies. These technologies contribute to enhancing sustainability and efficiency of social systems by interacting with physical infrastructures, and such systems integrating physical entities and cyber components, are referred as Cyber-Physical Systems (CPS). Many systems, such as energy, transportation, medical, and manufacturing systems, are considered as CPS [1], [2], and along with the innovation of Internet of Things (IoT), the importance of the systems will further increase. However, CPS are severely depending on computing and networking elements, and hence CPS have several opportunities for malicious third parties to inject attacks [3]–[5]. According to the solid interaction between cyber and physical space, the adversaries' action against CPS conduces tragic consequences to physical entities, not only cyber components. In other words, the cyber security of CPS is no longer restricted to the cyber domain. As a matter of fact, there exist disruptive cyber attacks targeting physical control system, which resulted in evil consequences to the physical components, e.g., the Stuxnet incident [6], the Maroochy water breach [7], and recent Ukrainian CyberAttacks [8]. For the sake of secure operations of CPS, therefore, we need to consider the cyber threat scenarios and strengthen security and resilience of them.

One major challenge to a secure operation of CPS is identifying the vulnerabilities due to malicious attacks and developing countermeasures and mitigations against them [9]–[18]. However, IoT devices have been increasing as well as cyber incidents and vulnerabilities are also increasing, and it is generally difficult to ensure the security of all devices and sensors. Additionally, in typical CPS such as energy systems, it is also difficult to immediately stop the operation even if they are subjected to malicious attacks. Thus, one another challenge is to operate CPS securely even if in the presence of malicious integrity attacks.

Toward this end, it is necessary to estimate the system state from the corrupted sensor measurements, and this problem is broadly referred as a secure state estimation problem. The paper [19] is pioneering work of the problem, and the authors derived necessary and sufficient conditions for the feasible estimation and control in linear systems and also provided an efficient state reconstruction algorithm based on compressive sensing literature. Shoukry and Tabuada [20] further refined this work and developed more efficient algorithm adopting an event-triggered technique. Chong *et al.* derived a new concept of the system observability in the presence of malicious attacks [21]. The papers [22]–[24] extended the problem into noisy systems, and the authors provided an ℓ_1-based attack-resilient state estimator with considering bounded noises and modeling errors. Lee *et al.* independently proposed an observer-based secure state estimator in the same environment [25]. In the paper [26], Nakahira and Mo developed a dynamic secure state estimator and provided an estimation-error bound on the ℓ_∞ norm. Our previous paper [27] presented a resilient estimator considering the reach set of the state. In the papers [28], [29], the authors developed computationally-efficient algorithms using Satisfiability Modulo Theory (SMT) paradigm. Further, some papers [30]–[32] discussed this problem in nonlinear systems.

This work was supported by JST CREST Grant Number JPMJCR15K2, Japan.

T. Shinohara and T. Namerikawa are with Department of System Design Engineering, Keio University, 3-14-1 Hiyoshi, Kohoku-ku, Yokohama, Kanagawa 223-8522, Japan. E-mail: takumis@nl.sd.keio.ac.jp and namerikawa@sd.keio.ac.jp

Z. Qu is with the Department of Electrical and Computer Engineering, University of Central Florida, Orlando, FL32816, USA. qu@ucf.edu

B. Motivation

Consider the following linear system subjected to malicious sensor attacks:

$$x(k+1) = Ax(k) + Bu(k), \tag{1}$$
$$y(k) = Cx(k) + y^a(k), \tag{2}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{m\times q}$, $C \in \mathbb{R}^{m\times n}$, $x(k) \in \mathbb{R}^n$ is the system state, $u(k) \in \mathbb{R}^q$ is the external input, $y(k) \in \mathbb{R}^m$ is the compromised sensor measurement, and $y^a(k) \in \mathbb{R}^m$ is the attack vector which is designed by the malicious adversary. Assume that the attack injection vector is sparse, and let us denote the number of nonzero entries of the vector by l. For the feasible state estimation under the sensor attacks, the following condition is well-known (see [19, Proposition 2], [20, Theorem 3.2], or [25, Lemma 3]):

Theorem 1: The system state can be reconstructed (possible with delay) if and only if $(A, C_{\mathcal{K}^{\complement}})$ is observable for every index set $\mathcal{K} \subset \{1,\dots,m\}$ with $|\mathcal{K}| = 2l$, where $\mathcal{K}^{\complement} \triangleq \{1,\dots,m\}\backslash\mathcal{K}$ is the complement set of $\mathcal{K}$ with respect to the sensor index $\{1,\dots,m\}$ and $C_{\mathcal{K}^{\complement}}$ is the matrix obtained from C by eliminating all the rows except those indexed by $\mathcal{K}^{\complement}$.

In other words, the state can be reconstructed if and only if the system remains observable after getting rid of any choice of $2l$ sensors. This condition obviously provides the maximum number of the compromised sensors: the attack on $m/2$ or more sensors precludes the state reconstruction since the observation matrix after removing $2l = m$ sensors will be the null matrix. Hence, the maximal number of the compromised sensors is equal to the half of the number of all sensors.

The above papers do not restrict the domain of the state to be reconstructed. The problem we are interested in is the secure state estimation problem when we have some information about the state. More precisely, can we reconstruct the state by using the *a priori* information even though more sensors are compromised? If possible, we can enhance the system resilience against sensor attacks using the information. In this paper, we especially focus on three prior information: the *sparsity information*, $(\alpha, \bar{n}_0)$*-sparsity information*, and *side information*, and, in each scenario, we derive that the system resilience is indeed reinforced by leveraging the information. Furthermore, we extend this analysis to noisy systems, and it is shown that the prior information suppresses the state-estimation error due to the measurement noise.

The reminder of this paper is organized as follows: In Section II, we describe the setup for the secure state estimation problem. Section III addresses the problem with the sparsity information, and we give a necessary and sufficient condition and a precise optimization problem for the secure state estimation with the information. Then, in Section IV, we extend this analysis into special sparsity scenario named $(\alpha, \bar{n}_0)$-sparsity condition, while Section V deals with the side information. Via numerical simulations resorting to a diffusion process in Section VI, we show that the prior information indeed enhances the system resilience. Section VII further extends this analysis to noisy systems, and it is presented the state-estimation error is suppressed by exploiting the prior information. Section VIII finally concludes this work.

Notation: For a set $\mathcal{I}$, $|\mathcal{I}|$ denotes the cardinality of the set. For a linear map A, $\ker A$ is denoted to the null space of A. We use $\mathbf{1}_n$ to indicate the n-dimensional column vector whose entries are 1. Given a vector $x \in \mathbb{R}^n$ and an index set $\mathcal{I} \subseteq \{1,\dots,n\}$, $x_{\mathcal{I}} \in \mathbb{R}^{|\mathcal{I}|}$ denotes the sub-vector formed by the entries indexed by $\mathcal{I}$. Similarly for a matrix $A \in \mathbb{R}^{m\times n}$ and an index set $\mathcal{J} \subseteq \{1,\dots,m\}$, $A_{\mathcal{J}} \in \mathbb{R}^{|\mathcal{J}|\times n}$ is the matrix obtained from A by eliminating all rows except those indexed by $\mathcal{J}$. Again for a vector $x \in \mathbb{R}^n$, the support of the vector, which is the set of the nonzero entries of x, is defined as

$$\operatorname{supp}(x) \triangleq \{i : x_i \neq 0\} \subseteq \{1,\dots,n\}.$$

The notion $\|x\|_{\ell_p}$ indicates the ℓ_p norm of a vector $x \in \mathbb{R}^n$ for some $p \geq 1$. Especially, we simply denote the ℓ_1 norm of the vector by $\|x\|_1$. Moreover, the ℓ_0 "norm" of the vector is denoted by $\|x\|_0$, which is defined as $\|x\|_0 \triangleq |\operatorname{supp}(x)|$, i.e., the number of the nonzero entries of the vector. We call a vector $x \in \mathbb{R}^n$ l-sparse if $\|x\|_0 \leq l$. Finally, the notation $\{x\}_0^{T-1}$ is used to denote a discrete-time finite-horizon sequence $\{x(0),\dots,x(T-1)\}$ for some $T \geq 1$, but we sometimes use $\{x\}$ ignoring the sub- and superscripts when there is no confusion.

II. Problem Formulation

In this paper, we first consider the secure state estimation problem in the noiseless linear time-invariant system which is modeled as (1) and (2). Since the external input is not affected by malicious sensor attacks and since the system is linear and known, we know from the principle of superposition that both $y(k)$ and $x(k)$ contain a zero-input response and a zero-initial-condition response. The latter can be explicitly computed based on $u(k)$, and hence we can assume in the subsequent analysis $u(k) \equiv 0$ for the purpose of determining the zero-input response. For the sake of readability, let us define the sensor index set as $\mathcal{S} \triangleq \{1,\dots,m\}$ and the state index set as $\mathcal{N} \triangleq \{1,\dots,n\}$. As aforementioned, the attack vector $y^a(k)$ is l-sparse, that is, $\|y^a(k)\|_0 \leq l,\ \forall k \in \mathbb{N}_0$. The subset of sensors the adversary can access to be assumed to be fixed over time, and is defined as $\mathcal{K} \subseteq \mathcal{S}$.

Our aim here is to recover the state $x(k-T+1)$ and to infer the attack vector sequence $\{y^a(k-T+1),\dots,y^a(k)\}$ from T (with $T \leq n$) sensor measurements $\{y(k-T+1),\dots,y(k)\}$. In this paper, without loss of generality, we focus on the reconstruction problem of the initial state $x(0)$ and the attack sequence $\{y^a\}$ from the sensor measurements $\{y\}$. For the problem, we assume that the defender has some knowledge of the initial state, which cannot be manipulated by the attacker. Thus, in what follows, we tackle the following problem.

Problem 1: For the system (1) and (2), given a potentially compromised measurement sequence $\{y\}$, recover the unique initial state $x(0)$ and attack sequence $\{y^a\}$ by leveraging available prior information.

We will provide necessary and sufficient conditions to this problem by deriving reconstructing conditions and convex-relaxation ones in three prior information. As prior information, Sections III, IV, and V focus on the sparsity information,

$(\alpha, \bar{n}_0)$-sparsity information, and side information of the state, respectively. The sparsity information indicates the number of the nonzero entries of the initial state while $(\alpha, \bar{n}_0)$-sparsity information can be applied when almost all the entries in the initial condition vector assume a specific value α. The side information, which can be regarded as an additional measurement not manipulated by the adversaries, indicates the physical description of the system initial state. We first consider the noiseless system, while the impact of measurement noises will be investigated in Section VII.

III. Secure State Estimation with Sparsity Information

In this section, we tackle the secure state estimation problem with sparsity information of the initial state. Denote the sparsity information by $\bar{n}_0$ such that $\|x(0)\|_0 \leq \bar{n}_0$. Note that if all entries in $x(0)$ are not 0, then $\bar{n}_0 = n$.

A. Necessary and Sufficient Condition for Estimation

We first give a necessary and sufficient condition for the secure state estimation under l sensor attacks with the sparsity information.

Theorem 2: For the system (1) and (2), suppose that potentially compromised T measurements $\{y\}$ are given. Any $\bar{n}_0$-sparse initial state and l-sparse attack sequence can be reconstructed from the measurements if and only if, for all $z \in \mathbb{R}^n \backslash \{0\}$ with $\|z\|_0 \leq 2\bar{n}_0$, the following holds:

$$\left|\operatorname{supp}(Cz) \cup \cdots \cup \operatorname{supp}\left(CA^{T-1}z\right)\right| > 2l. \tag{3}$$

Proof: First, for the sufficiency, we resort to a contradiction, that is, suppose that (3) satisfies, but all $\bar{n}_0$-sparse initial states and l-sparse attack sequences cannot be reconstructed. This means that there exist two initial states $x^1 \neq x^2$ with $\|x^1\|_0 \leq \bar{n}_0$ and $\|x^2\|_0 \leq \bar{n}_0$ and two attack scenarios $\{y^{a1}\}$ and $\{y^{a2}\}$ with $\operatorname{supp}\left(\{y^{a1}\}\right) \subseteq \mathcal{K}^1$, $|\mathcal{K}^1| \leq l$ and $\operatorname{supp}\left(\{y^{a2}\}\right) \subseteq \mathcal{K}^2$, $|\mathcal{K}^2| \leq l$ which lead to same output sequence $y(k) = CA^k x^1 + y^{a1}(k) = CA^k x^2 + y^{a2}(k)$[1]. Since the resulting outputs are same, it follows that

$$\left(y^{a1}(k) - y^{a2}(k)\right) = CA^k\left(x^2 - x^1\right), \ \forall k \in \mathbb{N}_0, \tag{4}$$

which is interpreted as that an initial state $x^2 - x^1 \neq 0$ generates an output sequence $y^{a1}(k) - y^{a2}(k)$. Note that this initial state yields $\|x^2 - x^1\|_0 \leq 2\bar{n}_0$ and the output satisfies $\operatorname{supp}\left(y^{a1}(k) - y^{a2}(k)\right) \subseteq \mathcal{K}^1 \cup \mathcal{K}^2$, $\forall k \in \mathbb{N}_0$. Thus, for a vector $z \triangleq x^2 - x^1$, which is $2\bar{n}_0$-sparse, we obtain $\operatorname{supp}\left(CA^k z\right) \subseteq \mathcal{K}^1 \cup \mathcal{K}^2$, $\forall k \in \mathbb{N}_0$, which shows that $|\operatorname{supp}\left(CA^k z\right)| \leq 2l$, $\forall k \in \mathbb{N}_0$, and thus (3) does not hold.

On the other hand, for the necessity, we again resort to a contradiction. Suppose that any $\bar{n}_0$-sparse initial state and any l-sparse attack sequence can be reconstructed, but (3) does not hold, which is equivalent to that there exists a vector $z \in \mathbb{R}^n \backslash \{0\}$ with $\|z\|_0 \leq 2\bar{n}_0$ satisfying

$$\left|\operatorname{supp}(Cz) \cup \cdots \cup \operatorname{supp}\left(CA^{T-1}z\right)\right| \leq 2l. \tag{5}$$

[1]Note that any two distinct initial states are distinguishable (i.e., these initial states do not generate same output sequence) if and only if every initial state can be reconstructed from T sensor measurements. For details, please refer to [33, Chapter 3] or other control theory literature.

In this proof, using two distinct $\bar{n}_0$-sparse vectors x^1 and x^2, let $z \triangleq x^1 - x^2$. Further, let $\mathcal{K}^1$ and $\mathcal{K}^2$ be two subsets of $\mathcal{S}$ with $|\mathcal{K}^1| \leq l$ and $|\mathcal{K}^2| \leq l$ such that $\mathcal{K}^1 \cup \mathcal{K}^2 = \operatorname{supp}(Cz) \cup \cdots \cup \operatorname{supp}\left(CA^{T-1}z\right)$. Note that $|\mathcal{K}^1 \cup \mathcal{K}^2| \leq 2l$. Consider the two attack scenarios are defined by $y^{a1}(k) = \left(-CA^k z\right)_i$, $i \in \mathcal{K}^1$, indicating the vector obtained from $-CA^k z$ by letting all the entries outside of $\mathcal{K}^1$ to 0, and similarly $y^{a2}(k) = \left(CA^k z\right)_i$, $i \in \mathcal{K}^2$. Then, it follows that

$$y(k) = CA^k x^1 + y^{a1}(k) = CA^k x^2 + y^{a2}(k), \ \forall k \in \mathbb{N}_0,$$

which obviously indicates two initial states and attack sequences correspond to the same output. Thus, two $\bar{n}_0$-sparse different initial states x^1, x^2 are not distinguishable, which implies a contradiction. ∎

The sparser the initial state is (or the smaller $\bar{n}_0$ is), the greater the left-hand side of (3) is, which implies that the sparse initial state enables the feasible secure state reconstruction even though more sensors are compromised. This indicates that the sparsity information of the state enhances the system resilience against sensor attacks. Note that if the initial state is not sparse (i.e., $\bar{n}_0 \geq n/2$), the necessary and sufficient condition of Theorem 2 coincides with the one of [19, Proposition 3.2]. Hence, in contrast, if the sparsity of initial state satisfies $\bar{n}_0 < n/2$, the state reconstruction might be achieved even if more sensors are compromised.

To illustrate the resilience reinforcement more closely, consider the case of $T = 1$. For a given initial output

$$y(0) = \underbrace{\begin{bmatrix} C & I \end{bmatrix}}_{\bar{C}} \underbrace{\begin{bmatrix} x(0) \\ y^a(0) \end{bmatrix}}_{\bar{x}} = \bar{C}\bar{x}, \quad \|\bar{x}\|_0 \leq \bar{n}_0 + l,$$

consider the problem to construct the vector $\bar{x} \in \mathbb{R}^{n+m}$. We here assume that C has full column rank (namely, the system is observable in a static sense). To this end, we first introduce the notion of the spark of a matrix [34], [35]:

Definition 1: For a matrix $A \in \mathbb{R}^{m \times n}$, the spark of the matrix is defined as

$$\operatorname{spark}(A) \triangleq \underset{z \in \ker A \backslash \{0\}}{\operatorname{minimize}} \|z\|_0. \tag{6}$$

In other words, the spark of A is the smallest number j such that there exists a set of j columns in A which are linear dependent.

Using this notion, the following proposition presents the feasible unique reconstruction of the solution.

Proposition 1 ([37, Corollary 1]): If $\|\bar{x}\|_0 < \operatorname{spark}(\bar{C})/2$, then $\bar{x}$ is the unique sparsest solution of $y(0) = \bar{C}\bar{x}$.

Now C has full column rank, and thus we have $\operatorname{spark}(\bar{C}) = m+1$. Therefore, by Proposition 1, if the following holds, then $\bar{x}$ obeying $\|\bar{x}\|_0 \leq \bar{n}_0 + l$ can be uniquely determined:

$$\bar{n}_0 + l < \frac{m+1}{2} \implies l < \frac{m+1}{2} - \bar{n}_0,$$

which indicates that a small $\bar{n}_0$ permits a large l. Thus, one can confirm that the sparse initial state can be recovered even if the adversary corrupts many sensors. The following simple example illustrates that the sparse initial state reinforces the system resilient against sensor attacks compared with a non-sparse initial state.

Example 1: Consider the case of $T = 1$. Then, the necessary and sufficient condition of Theorem 2 can be written as $\|Cz\|_0 > 2l,\ \forall z \in \mathbb{R}^n \backslash \{0\}$ with $\|z\|_0 \leq 2\bar{n}_0$. Let us assume that C is given as

$$C = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 & 2 \end{bmatrix}^\top. \tag{7}$$

If we have the knowledge that the initial state is sparse and satisfies $\|x(0)\|_0 \leq 1$, then it holds that, for all $z \in \mathbb{R}^n \backslash \{0\}$ with $\|z\|_0 \leq 2$, $\|Cz\|_0 > 3$, which implies that any 1-sparse initial state can be recovered under any 1 sensor is corrupted. On the other hand, if the initial state is not sparse (i.e., $\bar{n}_0 > 2$), then, for a vector $z = \begin{bmatrix} -1 & -1 & 1 \end{bmatrix}^\top$, we obtain $Cz = \begin{bmatrix} -2 & 0 & 0 & 0 & 1 \end{bmatrix}^\top$ and $\|Cz\|_0 = 2$, which indicates that any initial state cannot be reconstructed in the presence of 1-sparse attack. As a matter of fact, letting two distinct initial states be $x^1 = \begin{bmatrix} -1 & -1 & 1 \end{bmatrix}^\top$ and $x^2 = 0$ and two different 1-sparse attacks be $y^{a1} = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \end{bmatrix}^\top$ and $y^{a2} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \end{bmatrix}^\top$, it follows that

$$y = Cx^1 + y^{a1} = Cx^2 + y^{a2} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \end{bmatrix}^\top,$$

which implies two different attack scenarios explain same output, and thus these cannot be uniquely reconstructed.

B. Optimization Problem

In this subsection, we develop a concrete optimization problem to reconstruct the unique initial state and malicious attack sequence. For the system (1) and (2), the following holds regarding the collected outputs[2]:

$$Y = \Phi(x(0)) + Y^a, \tag{8}$$

where

$$\begin{aligned} Y &\triangleq \begin{bmatrix} y(0) & \cdots & y(T-1) \end{bmatrix} \in \mathbb{R}^{m \times T}, \\ Y^a &\triangleq \begin{bmatrix} y^a(0) & \cdots & y^a(T-1) \end{bmatrix} \in \mathbb{R}^{m \times T}, \\ \Phi(x) &: \mathbb{R}^n \to \mathbb{R}^{m \times T} \\ & \quad x \mapsto \begin{bmatrix} Cx & \cdots & CA^{T-1}x \end{bmatrix}. \end{aligned}$$

Based on the equation, let us consider the following optimization problem for some $p \geq 1$:

$$\begin{aligned} P_0^s : \underset{\hat{x} \in \mathbb{R}^n, \hat{Y}^a \in \mathbb{R}^{m \times T}}{\text{minimize}} \quad & \left\| \hat{Y}^a \right\|_{0/p} + \|\hat{x}\|_0 \\ \text{subject to} \quad & Y = \Phi(\hat{x}) + \hat{Y}^a, \end{aligned} \tag{9}$$

where $\hat{Y}^a \triangleq \begin{bmatrix} \hat{y}^a(0) & \cdots & \hat{y}^a(T-1) \end{bmatrix} \in \mathbb{R}^{m \times T}$ is an optimization variable pertaining to the attack sequence and, for a matrix $A \in \mathbb{R}^{m \times n}$, the ℓ_0/ℓ_p $(p \geq 1)$ mixed norm of the matrix is defined as follows [38]:

$$\|A\|_{0/p} \triangleq \sum_{i=1}^{m} \mathbb{I}\left(\|A_i\|_{\ell_p} \right),$$

[2]Though each matrix and the linear map depend on the time window T, we abbreviate it for the sake of legibility. Moreover, we sometimes abuse the notion of l-sparse for the matrix Y^a, meaning the attack sequence $\{y^a\}$ configuring this matrix is l-sparse.

where $\mathbb{I}(\cdot)$ is the indicator function, which is 0 when its argument is zero and is 1 otherwise. For the subsequent analysis, we further define the ℓ_1/ℓ_p $(p \geq 1)$ mixed norm of the matrix as

$$\|A\|_{1/p} \triangleq \sum_{i=1}^{m} \|A_i\|_{\ell_p}.$$

Thus, the optimization problem P_0^s can be observed as a reconstruction problem of the initial state and attack sequence, which explain the output matrix Y, minimizing the number of attacked sensors and nonzero entries of the initial state. Then, we obtain the following proposition for the feasible reconstruction.

Proposition 2: If, for all $z \in \mathbb{R}^n \backslash \{0\}$ with $\|z\|_0 \leq 2\bar{n}_0$, (3) holds, then any $\bar{n}_0$-sparse initial state and l-sparse attack matrix are the unique minimizers to P_0^s, i.e., for any $x(0) \in \mathbb{R}^n$ with $\|x(0)\|_0 \leq \bar{n}_0$ and $\{y^a\} \subset \mathbb{R}^m$ such that $\operatorname{supp}(\{y^a\}) \subseteq \mathcal{K}$ with $|\mathcal{K}| \leq l$, P_0^s can recover $x(0)$ and $\{y^a\}$.

Proof: We proceed by contradiction. Suppose that, for all $z \in \mathbb{R}^n \backslash \{0\}$ with $\|z\|_0 \leq 2\bar{n}_0$, (3) holds, but P_0^s cannot recover a unique $\bar{n}_0$-sparse initial state $x(0)$ and l-sparse attack sequence $\{y^a\}$. Let $(\hat{x}, \hat{Y}^a)$ be the solution of P_0^s, where $\hat{x} \neq x(0)$ and $\hat{Y}^a$ is a matrix concatenating an attack sequence $\{\hat{y}^a\} \neq \{y^a\}$ such that $\operatorname{supp}(\{\hat{y}^a\}) \subseteq \hat{\mathcal{K}}$. Note that, by the constraint condition of the P_0^s, we obtain $Y = \Phi(x(0)) + Y^a = \Phi(\hat{x}) + \hat{Y}^a$, with in addition, by the objective function, we also have $|\hat{\mathcal{K}}| + \|\hat{x}\|_0 \leq |\mathcal{K}| + \|x(0)\|_0 \leq l + \bar{n}_0$. Therefore, it is easy to see that two different sparse initial states and two different l-sparse attack sequences explain same output, which implies that these states and attack sequences cannot be reconstructed. This contradicts the condition of Theorem 2. ■

This proposition shows that the optimization problem P_0^s can recover the unique $\bar{n}_0$-sparse initial state and l-sparse attack sequence if (3) holds. However, this problem is, in general, NP-hard [39], and is not tractable. Thus, we next tackle the relaxation of this problem.

C. Relaxation to Convex Problem

Let us consider the following convex optimization problem for some $p \geq 1$:

$$\begin{aligned} P_{1/p}^s : \underset{\hat{x} \in \mathbb{R}^n, \hat{Y}^a \in \mathbb{R}^{m \times T}}{\text{minimize}} \quad & \left\| \hat{Y}^a \right\|_{1/p} + \|\hat{x}\|_1 \\ \text{subject to} \quad & Y = \Phi(\hat{x}) + \hat{Y}^a. \end{aligned} \tag{10}$$

Then, the following proposition provides a necessary and sufficient condition for the equivalence of the solutions between P_0^s and $P_{1/p}^s$.

Proposition 3: Suppose that, for all $z \in \mathbb{R}^n \backslash \{0\}$ with $\|z\|_0 \leq 2\bar{n}_0$, (3) holds. Letting $(x(0), Y^a)$, where $x(0)$ is $\bar{n}_0$-sparse and Y^a is a matrix concatenating an l-sparse attack sequence, be the solution of P_0^s and $(\hat{x}, \hat{Y}^a)$ be the one of $P_{1/p}^s$, $(x(0), Y^a) = (\hat{x}, \hat{Y}^a)$ if and only if, for all $\mathcal{K} \subset \mathcal{S}$ with

$|\mathcal{K}| = l$ and for all $\mathcal{N}_0 \subset \mathcal{N}$ with $|\mathcal{N}_0| = \bar{n}_0$, the following holds for all $x \in \mathbb{R}^n \backslash \{0\}$:

$$\sum_{i\in\mathcal{K}} \|(\Phi(x))_i\|_{\ell_p} + \sum_{j\in\mathcal{N}_0} |x_j| < \sum_{i\in\mathcal{K}^{\complement}} \|(\Phi(x))_i\|_{\ell_p} + \sum_{j\in\mathcal{N}_0^{\complement}} |x_j|\,, \tag{11}$$

where $\mathcal{K}^{\complement} \triangleq \mathcal{S}\backslash\mathcal{K}$ and $\mathcal{N}_0^{\complement} \triangleq \mathcal{N}\backslash\mathcal{N}_0$.

Proof: For the sufficiency, we resort to a contradiction. Suppose that, for all $\mathcal{K} \subset \mathcal{S}$ with $|\mathcal{K}| = l$ and for all $\mathcal{N}_0 \subset \mathcal{N}$ with $|\mathcal{N}_0| = \bar{n}_0$, (11) holds, but $(x(0), Y^a) \neq (\hat{x}, \hat{Y}^a)$, where $(x(0), Y^a)$ and $(\hat{x}, \hat{Y}^a)$ are, respectively, the solutions of P_0^s and $P_{1/p}^s$. By Proposition 2, we know that the solution of P_0^s is the unique one, which indicates that $P_{1/p}^s$ fails to recover the unique $\bar{n}_0$-sparse initial state $x(0)$ and l-sparse attack sequence $\{y^a\}$. Thus, by the objective function of $P_{1/p}^s$, two distinct solutions $(x(0), Y^a)$ and $(\hat{x}, \hat{Y}^a)$ follow

$$\left\|\hat{Y}^a\right\|_{1/p} + \|\hat{x}\|_1 \leq \|Y^a\|_{1/p} + \|x(0)\|_1\,. \tag{12}$$

Moreover, according to the constraint conditions of P_0^s and P_1^s, it follows that $Y = \Phi(x(0)) + Y^a = \Phi(\hat{x}) + \hat{Y}^a$. Here, let us define $\tilde{x} \triangleq \hat{x} - x(0) \neq 0$ following

$$\Phi(\tilde{x}) = \Phi(\hat{x}) - \Phi(x(0)) = Y^a - \hat{Y}^a. \tag{13}$$

Then, by the triangle inequality, we obtain

$$\begin{aligned}\sum_{i\in\mathcal{K}} \|(\Phi(\tilde{x}))_i\|_{\ell_p} + \sum_{j\in\mathcal{N}_0} |\tilde{x}_j| &= \sum_{i\in\mathcal{K}} \left\|Y_i^a - \hat{Y}_i^a\right\|_{\ell_p} + \sum_{j\in\mathcal{N}_0} |\tilde{x}_j| \\ &\geq \sum_{i\in\mathcal{K}} \left(\|Y_i^a\|_{\ell_p} - \left\|\hat{Y}_i^a\right\|_{\ell_p}\right) + \sum_{j\in\mathcal{N}_0} \left(|x_j(0)| - |\hat{x}_j|\right). \end{aligned} \tag{14}$$

Recalling Y^a is the unique attack matrix, the attack sequence configuring this matrix satisfies $\mathrm{supp}\,(\{y^a\}) \subseteq \mathcal{K}$. Thus, we have $\|Y^a\|_{1/p} = \sum_{i\in\mathcal{K}} \|Y_i^a\|_{\ell_p}$. Similarly, the unique initial state follows $\|x(0)\|_1 = \sum_{j\in\mathcal{N}_0} |x_j(0)|$. Hence, we obtain

$$\sum_{i\in\mathcal{K}} \|Y_i^a\|_{\ell_p} + \sum_{j\in\mathcal{N}_0} |x_j(0)| = \|Y^a\|_{1/p} + \|x(0)\|_1\,. \tag{15}$$

Then, (14) can be formulated as

$$\begin{aligned}&\sum_{i\in\mathcal{K}} \|(\Phi(\tilde{x}))_i\|_{\ell_p} + \sum_{j\in\mathcal{N}_0} |\tilde{x}_j| \\ &\geq \sum_{i\in\mathcal{K}} \left(\|Y_i^a\|_{\ell_p} - \left\|\hat{Y}_i^a\right\|_{\ell_p}\right) + \sum_{j\in\mathcal{N}_0} \left(|x_j(0)| - |\hat{x}_j|\right) \\ &\overset{(15)}{=} \|Y^a\|_{1/p} - \sum_{i\in\mathcal{K}} \left\|\hat{Y}_i^a\right\|_{\ell_p} + \|x(0)\|_1 - \sum_{j\in\mathcal{N}_0} |\hat{x}_j| \\ &\overset{(12)}{\geq} \left\|\hat{Y}^a\right\|_{1/p} - \sum_{i\in\mathcal{K}} \left\|\hat{Y}_i^a\right\|_{\ell_p} + \|\hat{x}\|_1 - \sum_{j\in\mathcal{N}_0} |\hat{x}_j| \\ &= \sum_{i\in\mathcal{K}^{\complement}} \left\|\hat{Y}_i^a\right\|_{\ell_p} + \sum_{j\in\mathcal{N}_0^{\complement}} |\hat{x}_j| = \sum_{i\in\mathcal{K}^{\complement}} \|(\Phi(\tilde{x}))_i\|_{\ell_p} + \sum_{j\in\mathcal{N}_0^{\complement}} |\tilde{x}_j|\,, \end{aligned}$$

where the last equality comes of the relation (13), $\mathrm{supp}\,(\{y^a\}) \subseteq \mathcal{K}$, and $\mathrm{supp}\,(x(0)) \subseteq \mathcal{N}_0$. This obviously contradicts (11).

For the necessity, we again consider a contradiction, namely, assume that $(x(0), Y^a) = (\hat{x}, \hat{Y}^a)$, but (11) does not hold, where, as with the sufficiency part, $(x(0), Y^a)$ and $(\hat{x}, \hat{Y}^a)$ are, respectively, the solutions of P_0^s and $P_{1/p}^s$. The vector $x(0)$ indicates the unique $\bar{n}_0$-sparse initial state and Y^a is the unique l-sparse attack matrix against the system. In other words, there exist an index set $\mathcal{K} \subset \mathcal{S}$ with $|\mathcal{K}| = l$, an index set $\mathcal{N}_0 \subset \mathcal{N}$ with $|\mathcal{N}_0| = \bar{n}_0$, and a vector $z \in \mathbb{R}^n \backslash \{0\}$ satisfying

$$\sum_{i\in\mathcal{K}} \|(\Phi(z))_i\|_{\ell_p} + \sum_{j\in\mathcal{N}_0} |z_j| \geq \sum_{i\in\mathcal{K}^{\complement}} \|(\Phi(z))_i\|_{\ell_p} + \sum_{j\in\mathcal{N}_0^{\complement}} |z_j|\,. \tag{16}$$

Using a nonzero initial state $x(0) \in \mathbb{R}^n \backslash \{0\}$ such that $\mathrm{supp}\,(x(0)) = \mathcal{N}_0$, let an attack matrix be $Y^a = -(\Phi(x(0)))_i,\ i \in \mathcal{K}$. Note that the attack sequence configuring this matrix satisfies $\mathrm{supp}\,(\{y^a\}) = \mathcal{K}$. The resulting output matrix is given by $Y = \Phi(x(0)) + Y^a = (\Phi(x(0)))_i,\ i \in \mathcal{K}^{\complement}$. In contrast, considering another initial state $\hat{x} = 0$ and attack matrix $\hat{Y}^a = (\Phi(x(0)))_i,\ i \in \mathcal{K}^{\complement}$, we have same output matrix $Y = \Phi(\hat{x}) + \hat{Y}^a = (\Phi(x(0)))_i,\ i \in \mathcal{K}^{\complement}$. By the definitions of Y^a and $x(0)$, it follows that

$$\begin{aligned}\|Y^a\|_{1/p} + \|x(0)\|_1 &= \sum_{i\in\mathcal{K}} \|Y_i^a\|_{\ell_p} + \sum_{j\in\mathcal{N}_0} |x_j(0)| \\ &= \sum_{i\in\mathcal{K}} \|(\Phi(x(0)))_i\|_{\ell_p} + \sum_{j\in\mathcal{N}_0} |x_j(0)| \\ &\overset{(16)}{\geq} \sum_{i\in\mathcal{K}^{\complement}} \|(\Phi(x(0)))_i\|_{\ell_p} + \sum_{j\in\mathcal{N}_0^{\complement}} |x_j(0)| \\ &= \left\|\hat{Y}^a\right\|_{1/p} + \|\hat{x}\|_1\,, \end{aligned} \tag{17}$$

where we use the fact that $\sum_{j\in\mathcal{N}_0^{\complement}} |x_j(0)| = \|\hat{x}\|_1 = 0$. This indicates that $P_{1/p}^s$ fails to reconstruct the unique initial state $x(0)$ and attack matrix Y^a, i.e., $(x(0), Y^a) \neq (\hat{x}, \hat{Y}^a)$, which implies a contradiction. ■

This proposition says that one can reconstruct any $\bar{n}_0$-sparse initial state and l-sparse attack sequence using a tractable convex optimization problem $P_{1/p}^s$ provided that the system satisfies (11). The condition (11) relates to the *nullspace property* in the field of compressed sensing and the calculation of this conditions is known to be NP-hard, in general [40]. Thus, there is, unfortunately, no known efficient way to calculate this condition. However, we note that this paper tackles the resilience reinforcement with *a priori* information, and as given in the following remark, one can realize that the prior information indeed enhances the system resilience when compared with the conventional result in terms of practical computation as well.

Remark 1: In the case when we do not consider the sparsity information, the condition of (11) is given as the following [19, Proposition 6]:

$$\sum_{i\in\mathcal{K}} \|(\Phi(x))_i\|_{\ell_p} < \sum_{i\in\mathcal{K}^{\complement}} \|(\Phi(x))_i\|_{\ell_p}\,.$$

We observe that, broadly speaking, this relation satisfies when l is small enough, namely the number of compromised sensors

is low, but the larger l is not permitted. On the other hand, in (11) considering the sparsity, the increase of l can be compensated by the increasing of the sparsity (or the decreasing of $\bar{n}_0$). Hence, also from the viewpoint of calculation, the sparsity helps the state reconstruction under the malicious sensor attacks. However, we likewise recognize that non-sparse initial states rather deteriorate the condition. Hence, when we know that the initial state is not sparse enough, the optimization problem without considering sparsity information (which is given as $\mathcal{D}_{1/p}$ in Section VI) should be utilized.

Remark 2: The initial state is mostly sparse in large and complex systems and the sparse information can be given in advance. A typical example is diffusion mechanisms such as the virus spreading in computer networks [41] or data diffusion process in networks [42]. It is well known that these systems have sparse initial states. However, these systems are laid on ubiquitous and complex networks, and thus adversarial attacks easily can corrupt the processes. Therefore, the analysis in this section helps the complex network systems admitting to sparse initial states to enhance the resilience against the malicious interruption.

Remark 3: As a final remark, it should be noted that there exists a trade-off between the system resilience and operational performance relying on the time window T. Choosing a large T indicates one can exploit many sensor measurements for the estimation. However, a large T generates a delay of the estimation which possibly deteriorates the operational performance. Therefore, we ought to select a suitable T satisfying the state reconstruction condition for the anticipated l. If the defender would like to stiffen the system resilience as far as possible, T is sufficient to be equal to the number of states n, by the Cayley-Hamilton theorem.

IV. Secure State Estimation with $(\alpha, \bar{n}_0)$-Sparsity Information

In the previous section, we addressed that the sparsity information enhances the system resilience. We further derived a convex optimization problem which recovers any $\bar{n}_0$-sparse initial state and l-sparse attack sequence under the condition (11). This section extends this sparsity information into a special sparsity case, i.e., the most entries of the state are an arbitrary value, not only 0, as follows.

Definition 2 ($(\alpha, \bar{n}_0)$-sparse vector): We call a vector $x \in \mathbb{R}^n$ $(\alpha, \bar{n}_0)$-sparse if the following holds for some $\alpha \in \mathbb{R}$:

$$\|x - \alpha \mathbf{1}_n\|_0 \leq \bar{n}_0. \tag{18}$$

In other words, if most entries of a vector x are α and the number of otherwise entries is $\bar{n}_0$, then the vector is called $(\alpha, \bar{n}_0)$-sparse. In general, however, note that the defender does not have the knowledge of α.

A. Necessary and Sufficient Condition for Estimation

In analogy with the previous section, we first give a necessary and sufficient condition for the secure state estimation in the presence of l sensor attacks with the $(\alpha, \bar{n}_0)$-sparsity information.

Theorem 3: For the system (1) and (2), suppose that potentially compromised T measurements $\{y\}$ are given. For some $\alpha \in \mathbb{R}$, any $(\alpha, \bar{n}_0)$-sparse initial state and l-sparse attack sequence can be reconstructed from the measurements if and only if, for all $z \in \mathbb{R}^n \backslash \{0\}$ with $\|z\|_0 \leq 2\bar{n}_0$, (3) holds.

Proof: In analogy with Theorem 2, for the sufficiency, we resort to a contradiction, that is, suppose that (3) satisfies, but all $(\alpha, \bar{n}_0)$-sparse initial states and attack sequences cannot be reconstructed. This means that there exist two distinct $(\alpha, \bar{n}_0)$-sparse initial states $x^1 \neq x^2$ and two attack scenarios $\{y^{a1}\}$ and $\{y^{a2}\}$ with $\mathrm{supp}\left(\{y^{a1}\}\right) \subseteq \mathcal{K}^1$, $|\mathcal{K}^1| \leq l$ and $\mathrm{supp}\left(\{y^{a2}\}\right) \subseteq \mathcal{K}^2$, $|\mathcal{K}^2| \leq l$ which lead to same output sequence $y(k) = CA^k x^1 + y^{a1}(k) = CA^k x^2 + y^{a2}(k)$. Hence, as with Theorem 2, we see that an initial state $x^2 - x^1 \neq 0$ generates an output sequence $y^{a1}(k) - y^{a2}(k)$. Note that this initial state satisfies $\|x^2 - x^1\|_0 \leq 2\bar{n}_0$ (this is because, for an index i, if it satisfies that $x_i^1 = x_i^2 = \alpha$, then we have $x_i^2 - x_i^1 = 0$) and the output satisfies $\mathrm{supp}\left(y^{a1}(k) - y^{a2}(k)\right) \subseteq \mathcal{K}^1 \cup \mathcal{K}^2$, $\forall k \in \mathbb{N}_0$. Thus, for a vector $z \triangleq x^2 - x^1$ which is $2\bar{n}_0$-sparse, we obtain $\mathrm{supp}\left(CA^k z\right) \subseteq \mathcal{K}^1 \cup \mathcal{K}^2$, $\forall k \in \mathbb{N}_0$, which shows that $|\mathrm{supp}\left(CA^k z\right)| \leq 2l$, $\forall k \in \mathbb{N}_0$, and thus (3) does not hold.

For the necessity, then, we again resort to a contradiction. Suppose that any $(\alpha, \bar{n}_0)$-sparse initial state and any l-sparse attack sequence can be reconstructed, but (3) does not hold, which is equivalent to that there exists a vector $z \in \mathbb{R}^n \backslash \{0\}$ with $\|z\|_0 \leq 2\bar{n}_0$ satisfying (5). As with Theorem 2, let us define $z \triangleq x^1 - x^2 \neq 0$, where x^1 and x^2 are $\bar{n}_0$-sparse. Note that these vectors are $(0, \bar{n}_0)$-sparse, and thus the rest of this proof is same as the one of Theorem 2. ∎

One can observe that even if the most entries of the initial state take the same value, not 0 but some $\alpha \in \mathbb{R}$, the condition for the state reconstruction is not altered. Thus, if the most entries of the initial state are the same value, the system resilience against sensor attacks is enhanced.

B. Optimization Problem

As is the case with the previous section, then, we construct an optimization problem to reconstruct the $(\alpha, \bar{n}_0)$-sparse initial state and malicious attack sequence. To this end, let us consider the following optimization problem for some $p \geq 1$:

$$\begin{aligned} P_0^\alpha : \underset{\substack{\hat{x} \in \mathbb{R}^n, \hat{\alpha} \in \mathbb{R}, \\ \hat{Y}^a \in \mathbb{R}^{m \times T}}}{\text{minimize}} \quad & \left\| \hat{Y}^a \right\|_{0/p} + \|\hat{x} - \hat{\alpha} \mathbf{1}_n\|_0 \\ \text{subject to} \quad & Y = \Phi(\hat{x}) + \hat{Y}^a. \end{aligned} \tag{19}$$

One difference from the problem P_0^s is to take α into account since the defender does not have the knowledge of α. Then, we obtain the following proposition pertaining to P_0^α.

Proposition 4: If, for all $z \in \mathbb{R}^n \backslash \{0\}$ with $\|z\|_0 \leq 2\bar{n}_0$, (3) holds, then any $(\alpha, \bar{n}_0)$-sparse initial state and l-sparse attack matrix are the unique minimizers to P_0^α, i.e., for any $x(0) \in \mathbb{R}^n$ with $\|x(0) - \alpha \mathbf{1}_n\|_0 \leq \bar{n}_0$ and for any $\{y^a\} \subset \mathbb{R}^m$ such that $\mathrm{supp}\left(\{y^a\}\right) \subseteq \mathcal{K}$ with $|\mathcal{K}| \leq l$, P_0^α can recover $x(0)$, α, and $\{y^a\}$.

Proof: With the aim of contradiction, suppose that, for all $z \in \mathbb{R}^n \backslash \{0\}$ with $\|z\|_0 \leq 2\bar{n}_0$, (3) holds, but P_0^α cannot recover a unique $(\alpha, \bar{n}_0)$-sparse initial state $x(0)$ and l-sparse attack sequence $\{y^a\}$. Let $(\hat{x}, \hat{\alpha}, \hat{Y}^a)$ be the solution of P_0^α,

where $\hat{x} \neq x(0)$ with $\|\hat{x} - \hat{\alpha}\mathbf{1}_n\|_0 \leq \bar{n}_0$ for some $\hat{\alpha} \in \mathbb{R}$ and $\hat{Y}^a$ is a matrix concatenating an attack sequence $\{\hat{y}^a\} \neq \{y^a\}$ such that $\mathrm{supp}\,(\{\hat{y}^a\}) \subseteq \hat{\mathcal{K}}$. Note that, by the constraint condition of the P_0^α, we have $Y = \Phi(x(0)) + Y^a = \Phi(\hat{x}) + \hat{Y}^a$, with in addition, by the objective function, we also have $|\hat{\mathcal{K}}| + \|\hat{x} - \hat{\alpha}\mathbf{1}_n\|_0 \leq |\mathcal{K}| + \|x(0) - \alpha\mathbf{1}_n\|_0 \leq l + \bar{n}_0$. Therefore, it is easy to see that two different $(\alpha, \bar{n}_0)$-sparse initial states and two different l-sparse attack sequences explain same output, which implies that these states and attack sequences cannot be reconstructed. This contradicts the condition of Theorem 3. ■

C. Relaxation to Convex Problem

As with the previous section, we then tackle to relax the problem P_0^α into a tractable one. Consider the following convex optimization problem for some $p \geq 1$:

$$P_{1/p}^\alpha : \underset{\substack{\hat{x}\in\mathbb{R}^n, \hat{\alpha}\in\mathbb{R}, \\ \hat{Y}^a \in \mathbb{R}^{m\times T}}}{\text{minimize}} \quad \left\|\hat{Y}^a\right\|_{1/p} + \|\hat{x} - \hat{\alpha}\mathbf{1}_n\|_1 \\ \text{subject to} \quad Y = \Phi(\hat{x}) + \hat{Y}^a. \tag{20}$$

Then, the following proposition derives a necessary and sufficient condition for the equivalence of the solutions between P_0^α and $P_{1/p}^\alpha$.

Proposition 5: Suppose that, for all $z \in \mathbb{R}^n\backslash\{0\}$ with $\|z\|_0 \leq 2\bar{n}_0$, (3) holds. Letting $(x(0), \alpha, Y^a)$, where $x(0)$ is $(\alpha, \bar{n}_0)$-sparse for $\alpha \in \mathbb{R}$ and Y^a is a matrix concatenating an l-sparse attack sequence, be the solution of P_0^α and $(\hat{x}, \hat{\alpha}, \hat{Y}^a)$ be the one of $P_{1/p}^\alpha$, $(x(0), \alpha, Y^a) = (\hat{x}, \hat{\alpha}, \hat{Y}^a)$ if and only if, for all $\mathcal{K} \subset \mathcal{S}$ with $|\mathcal{K}| = l$ and for all $\mathcal{N}_0 \subset \mathcal{N}$ with $|\mathcal{N}_0| = \bar{n}_0$, the following holds for all $x \in \mathbb{R}^n\backslash\{0\}$ and for all $\beta \in \mathbb{R}$:

$$\sum_{i\in\mathcal{K}} \|(\Phi(x))_i\|_{\ell_p} + \sum_{j\in\mathcal{N}_0} |x_j - \beta| \\ < \sum_{i\in\mathcal{K}^\complement} \|(\Phi(x))_i\|_{\ell_p} + \sum_{j\in\mathcal{N}_0^\complement} |x_j - \beta|. \tag{21}$$

Proof: As well Proposition 3, for the sufficiency, we resort to a contradiction. Suppose that, for all $\mathcal{K} \subset \mathcal{S}$ with $|\mathcal{K}| = l$ and for all $\mathcal{N}_0 \subset \mathcal{N}$ with $|\mathcal{N}_0| = \bar{n}_0$, (21) holds, but $(x(0), \alpha, Y^a) \neq (\hat{x}, \hat{\alpha}, \hat{Y}^a)$, where $(x(0), \alpha, Y^a)$ and $(\hat{x}, \hat{\alpha}, \hat{Y}^a)$ are, respectively, the solutions of P_0^α and $P_{1/p}^\alpha$. By Proposition 4, we now know that the solution of P_0^α is the unique one, which indicates that $P_{1/p}^\alpha$ fails to recover the unique $(\alpha, \bar{n}_0)$-sparse initial state $x(0)$ and l-sparse attack sequence. Thus, two different solutions $(x(0), \alpha, Y^a)$ and $(\hat{x}, \hat{\alpha}, \hat{Y}^a)$ follow

$$\left\|\hat{Y}^a\right\|_{1/p} + \|\hat{x} - \hat{\alpha}\mathbf{1}_n\|_1 \leq \|Y^a\|_{1/p} + \|x(0) - \alpha\mathbf{1}_n\|_1. \tag{22}$$

Moreover, according to the constraint conditions of P_0^α and $P_{1/p}^\alpha$, we have $Y = \Phi(x(0)) + Y^a = \Phi(\hat{x}) + \hat{Y}^a$. Here, let us define $\tilde{\alpha} \triangleq \hat{\alpha} - \alpha$ and $\tilde{x} \triangleq \hat{x} - x(0) \neq 0$ following (13). Then, by the triangle inequality, we obtain

$$\sum_{i\in\mathcal{K}} \|(\Phi(\tilde{x}))_i\|_{\ell_p} + \sum_{j\in\mathcal{N}_0} |\tilde{x}_j - \tilde{\alpha}| \\ \geq \sum_{i\in\mathcal{K}} \left(\|Y_i^a\|_{\ell_p} - \left\|\hat{Y}_i^a\right\|_{\ell_p} \right) + \sum_{j\in\mathcal{N}_0} \left(|x_j(0) - \alpha| - |\hat{x}_j - \hat{\alpha}|\right) \tag{23}$$

As indicated in the proof of Proposition 3, we have $\|Y^a\|_{1/p} = \sum_{i\in\mathcal{K}} \|Y_i^a\|_{\ell_p}$. Similarly, regarding the unique initial state $x(0)$ and α, they satisfy $\|x(0) - \alpha\mathbf{1}_n\|_1 = \sum_{j\in\mathcal{N}_0} |x_j(0) - \alpha|$. Hence, by (22), it follows that

$$\sum_{i\in\mathcal{K}} \|Y_i^a\|_{\ell_p} + \sum_{j\in\mathcal{N}_0} |x_j(0) - \alpha| = \|Y^a\|_{1/p} + \|x(0) - \alpha\mathbf{1}_n\|_1 \\ \geq \left\|\hat{Y}^a\right\|_{1/p} + \|\hat{x} - \hat{\alpha}\mathbf{1}_n\|_1. \tag{24}$$

Then, by (23), we have

$$\begin{aligned} &\sum_{i\in\mathcal{K}} \|(\Phi(\tilde{x}))_i\|_{\ell_p} + \sum_{j\in\mathcal{N}_0} |\tilde{x}_j - \tilde{\alpha}| \\ &\overset{(24)}{\geq} \left\|\hat{Y}^a\right\|_{1/p} - \sum_{i\in\mathcal{K}} \left\|\hat{Y}_i^a\right\|_{\ell_p} + \|\hat{x} - \hat{\alpha}\mathbf{1}_n\|_1 - \sum_{j\in\mathcal{N}_0} |\hat{x}_j - \hat{\alpha}| \\ &= \sum_{i\in\mathcal{K}^\complement} \left\|\hat{Y}_i^a\right\|_{\ell_p} + \sum_{j\in\mathcal{N}_0^\complement} |\hat{x}_j - \hat{\alpha}| \\ &= \sum_{i\in\mathcal{K}^\complement} \|(\Phi(\tilde{x}))_i\|_{\ell_p} + \sum_{j\in\mathcal{N}_0^\complement} |\tilde{x}_j - \tilde{\alpha}|, \end{aligned}$$

where the last equality is caused by the relation (13), $\mathrm{supp}\,(\{y^a\}) \subseteq \mathcal{K}$, and $\mathrm{supp}\,(x(0) - \alpha\mathbf{1}_n) \subseteq \mathcal{N}_0$. This obviously contradicts (21).

For the necessity, we again consider a contradiction, namely, suppose that $(x(0), \alpha, Y^a) = (\hat{x}, \hat{\alpha}, \hat{Y}^a)$, but (21) does not hold, where, in analogy with the previous sufficiency part, $(x(0), \alpha, Y^a)$ and $(\hat{x}, \hat{\alpha}, \hat{Y}^a)$ are, respectively, the solutions of P_0^α and $P_{1/p}^\alpha$, where $x(0)$ is the unique $(\alpha, \bar{n}_0)$-sparse initial state and Y^a is the l-sparse attack matrix. In other words, there exist an index set $\mathcal{K} \subset \mathcal{S}$ with $|\mathcal{K}| = l$, an index set $\mathcal{N}_0 \subset \mathcal{N}$ with $|\mathcal{N}_0| = \bar{n}_0$, a vector $x \in \mathbb{R}^n\backslash\{0\}$, and a scalar $\beta \in \mathbb{R}$ such that

$$\sum_{i\in\mathcal{K}} \|(\Phi(x))_i\|_{\ell_p} + \sum_{j\in\mathcal{N}_0} |x_j - \beta| \\ \geq \sum_{i\in\mathcal{K}^\complement} \|(\Phi(x))_i\|_{\ell_p} + \sum_{j\in\mathcal{N}_0^\complement} |x_j - \beta|. \tag{25}$$

Using a nonzero $(\alpha, \bar{n}_0)$-sparse initial state $x(0) \in \mathbb{R}^n\backslash\{0\}$ such that $\mathrm{supp}\,(x(0) - \alpha\mathbf{1}_n) = \mathcal{N}_0$ for some $\alpha \in \mathbb{R}$, let an attack matrix be $Y^a = -(\Phi(x(0)))_i$, $i \in \mathcal{K}$. The resulting output matrix is given by $Y = \Phi(x(0)) + Y^a = (\Phi(x(0)))_i$, $i \in \mathcal{K}^\complement$. In contrast, considering another initial state $\hat{x} = 0$ and another attack matrix $\hat{Y}^a = (\Phi(x(0)))_i$, $i \in \mathcal{K}^\complement$, we have

same output matrix $Y = \Phi(\hat{x}) = \hat{Y}^a = (\Phi(x(0)))_i,\ i \in \mathcal{K}^{\complement}$. Defining $\hat{\alpha} \triangleq 0$, by the definitions of Y^a and $x(0)$, we have

$$\begin{aligned}
&\|Y^a\|_{1/p} + \|x(0) - \alpha \mathbf{1}_n\|_1 \\
&= \sum_{i\in\mathcal{K}} \|(\Phi(x(0)))_i\|_{\ell_p} + \sum_{j\in\mathcal{N}_0} |x_j(0) - \alpha|\,, \\
&\overset{(25)}{\geq} \sum_{i\in\mathcal{K}^{\complement}} \|(\Phi(x(0)))_i\|_{\ell_p} + \sum_{j\in\mathcal{N}_0^{\complement}} |x_j(0) - \alpha| \\
&= \left\|\hat{Y}^a\right\|_{1/p} + \|\hat{x} - \hat{\alpha}\mathbf{1}_n\|_1\,, \qquad (26)
\end{aligned}$$

where we use the fact that $\sum_{j\in\mathcal{N}_0^{\complement}} |x_j(0) - \alpha| = \|\hat{x} - \hat{\alpha}\mathbf{1}_n\|_1 = 0$. This inequality indicates that $P^{\alpha}_{1/p}$ fails to reconstruct the unique $(\alpha, \bar{n}_0)$-sparse initial state $x(0)$, α, and attack matrix Y^a, i.e., $(x(0), \alpha, Y^a) \neq (\hat{x}, \hat{\alpha}, \hat{Y}^a)$, which implies a contradiction. ∎

This proposition indicates that any $(\alpha, \bar{n}_0)$-sparse initial state and l-sparse attack sequence can be reconstructed using a tractable convex optimization problem $P^{\alpha}_{1/p}$ if the system satisfies (21). Note that, unlike the previous estimator $P^{s}_{1/p}$, one needs to recover α which configures the initial state. As aforementioned, the state reconstruction conditions for sparse and $(\alpha, \bar{n}_0)$-sparse initial state are same, but in practice, the recovery rate of the $(\alpha, \bar{n}_0)$-sparse initial state might be inferior to the sparse one.

V. Secure State Estimation with Side Information

As third prior information, we consider side information of the initial state which is given as follows [43]:

$$\psi = \Omega x(0), \qquad (27)$$

where $\psi \in \mathbb{R}^r$ is the side information and $\Omega \in \mathbb{R}^{r\times n}$ is called side information matrix. If Ω has full column rank, one can uniquely reconstruct the initial state using the matrix and ψ, and therefore, without loss of generality, we assume here that Ω has full row rank and is not invertible. The case when $r = 0$ (i.e., Ω is the null matrix) is equivalent to the case of no prior information of the initial state. As indicated in [43], the side information ψ implies the knowledge of the initial state from the physical attribution of the system and cannot be manipulated by malicious third parties.

A. Necessary and Sufficient Condition for Estimation

As with the discussion so far, we first give a necessary and sufficient condition for the secure state estimation under l sensor attacks with the side information of the initial state.

Theorem 4: For the system (1) and (2), suppose that potentially compromised T measurements $\{y\}$ are given. Additionally, suppose that the side information ψ and Ω are given. Any initial state satisfying (27) and l-sparse attack sequence can be reconstructed from the measurements if and only if, for all $z \in \mathbb{R}^n\backslash\{0\}$, the following holds:

$$\left|\operatorname{supp}\left(\begin{bmatrix} C \\ \Omega \end{bmatrix} z\right) \cup \cdots \cup \operatorname{supp}\left(\begin{bmatrix} CA^{T-1} \\ \Omega \end{bmatrix} z\right)\right| > 2l. \qquad (28)$$

Proof: In analogy with previous theorems, for the sufficiency, we resort to a contradiction, that is, suppose that (28) satisfies, but all initial states satisfying (27) and l-sparse attack sequences cannot be reconstructed. This means that there exist two initial states $x^1 \neq x^2$ and two attack scenarios $\{y^{a1}\}$ and $\{y^{a2}\}$ with $\operatorname{supp}(\{y^{a1}\}) \subseteq \mathcal{K}^1,\ |\mathcal{K}^1| \leq l$ and $\operatorname{supp}(\{y^{a2}\}) \subseteq \mathcal{K}^2,\ |\mathcal{K}^2| \leq l$ which lead to same output sequence $y(k) = CA^k x^1 + y^{a1}(k) = CA^k x^2 + y^{a2}(k)$. Hence, as is the case with Theorem 2, we see that an initial state $x^2 - x^1 \neq 0$ generates an output sequence $y^{a1}(k) - y^{a2}(k)$. Moreover, $x^2 - x^1$ yields $\Omega(x^2 - x^1) = 0$, and thus it follows that

$$\begin{bmatrix} y^{a1}(k) - y^{a2}(k) \\ 0 \end{bmatrix} = \begin{bmatrix} CA^k \\ \Omega \end{bmatrix} (x^2 - x^1)\,,\ \forall k \in \mathbb{N}_0.$$

Therefore, for a given nonzero vector $z \triangleq x^2 - x^1$, we obtain

$$\operatorname{supp}\left(\begin{bmatrix} CA^k \\ \Omega \end{bmatrix} z\right) \subseteq (\mathcal{K}^1 \cup \mathcal{K}^2) \times \{1, \ldots, r\},\ \forall k \in \mathbb{N}_0,$$

which implies, due to $|\mathcal{K}^1| \leq l$, $|\mathcal{K}^2| \leq l$, and the fact that $\Omega z = 0$,

$$\left|\operatorname{supp}\left(\begin{bmatrix} CA^k \\ \Omega \end{bmatrix} z\right)\right| \leq 2l.$$

This obviously contradicts (28).

For the necessity, then, we again resort to a contradiction. Suppose that any initial state and any l-sparse attack sequence can be reconstructed, but (28) does not hold, which is equivalent to that there exists a vector $z \in \mathbb{R}^n\backslash\{0\}$ satisfying

$$\left|\operatorname{supp}\left(\begin{bmatrix} C \\ \Omega \end{bmatrix} z\right) \cup \cdots \cup \operatorname{supp}\left(\begin{bmatrix} CA^{T-1} \\ \Omega \end{bmatrix} z\right)\right| \leq 2l.$$

In this proof, split z into two vectors x^1 and x^2 so that $z \triangleq x^1 - x^2$, where $x^1 \in \ker\Omega$ and $x^2 \in \ker\Omega$. Let $\mathcal{K}^1$ and $\mathcal{K}^2$ be two subsets of $\mathcal{S}$ with $|\mathcal{K}^1| \leq l$ and $|\mathcal{K}^2| \leq l$ such that

$$\mathcal{K}^1 \cup \mathcal{K}^2 = \operatorname{supp}\left(\begin{bmatrix} C \\ \Omega \end{bmatrix} z\right) \cup \cdots \cup \operatorname{supp}\left(\begin{bmatrix} CA^{T-1} \\ \Omega \end{bmatrix} z\right).$$

By the definition, it yields that $\operatorname{supp}(\Omega z) = \emptyset$, and thus $\mathcal{K}^1 \cup \mathcal{K}^2$ can be reformulated as

$$\mathcal{K}^1 \cup \mathcal{K}^2 = \operatorname{supp}(Cz) \cup \cdots \cup \operatorname{supp}\left(CA^{T-1}z\right).$$

In analogy with Theorem 2, letting two attack sequences be $y^{a1}(k) = \left(-CA^k z\right)_i,\ i \in \mathcal{K}^1$ and $y^{a2}(k) = \left(CA^k z\right)_i,\ i \in \mathcal{K}^2$, we have $CA^k x^1 + y^{a1}(k) = CA^k x^2 + y^{a2}(k),\ \forall k \in \mathbb{N}_0$, which implies two distinct initial state x^1 and x^2 are distinguishable. ∎

This necessary and sufficient condition obviously shows the resilient reinforcement in secure state estimation, namely, if we have much side information, then the left-hand side of (28) can be greater than m. This implies that even though more than half of all sensors are compromised, the initial state can be recovered using this information. The following simple example shows the resilience reinforcement with the side information.

Example 2: As in Example 1, consider $T = 1$ and C is given as (7). We know, unless the initial state is sparse, the system is vulnerable for some 1-sparse attack. By contrast, if

we obtain the side information matrix Ω such that $\Omega x \neq 0$, where $x = [\ -1 \quad -1 \quad 1\ ]^\top$, it follows that

$$\left|\mathrm{supp}\left(\begin{bmatrix} C \\ \Omega \end{bmatrix} z\right)\right| > 2,\ \forall z \in \mathbb{R}^n \backslash \{0\}, \tag{29}$$

which indicates any initial state can be recovered under any 1-sparse attack.

B. Optimization Problem

Next, we construct an optimization problem to reconstruct the initial state and malicious attack sequence resorting to the side information. Let us consider the following optimization problem taking the side information as an equality constraint into account:

$$\begin{aligned} P_0^i : \underset{\hat{x}\in\mathbb{R}^n, \hat{Y}^a \in \mathbb{R}^{m\times T}}{\text{minimize}} \quad & \left\|\hat{Y}^a\right\|_{0/p} \\ \text{subject to} \quad & Y = \Phi(\hat{x}) + \hat{Y}^a, \\ & \psi = \Omega \hat{x}. \end{aligned} \tag{30}$$

Then, the following proposition is given.

Proposition 6: Suppose that the side information ψ and Ω are given. If, for all $z \in \mathbb{R}^n\backslash\{0\}$, (28) holds, then any initial state satisfying (27) and l-sparse attack matrix are the unique minimizers to P_0^i, i.e., for any initial state $x(0) \in \mathbb{R}^n$ obeying (27) and for any $\{y^a\} \subset \mathbb{R}^m$ such that $\mathrm{supp}\,(\{y^a\}) \subseteq \mathcal{K}$ with $|\mathcal{K}| \leq l$, P_0^i can recover $x(0)$ and $\{y^a\}$.

Proof: With the aim of contradiction, suppose that, for all $z \in \mathbb{R}^n\backslash\{0\}$, (28) holds, but P_0^i cannot recover a unique initial state $x(0)$ satisfying (27) and l-sparse attack sequence $\{y^a\}$. Let $(\hat{x}, \hat{Y}^a)$ be the solution of P_0^i, where $\hat{x}$ satisfies (27) but $\hat{x} \neq x(0)$ and $\hat{Y}^a$ is a matrix concatenating an attack sequence $\{\hat{y}^a\} \neq \{y^a\}$ such that $\mathrm{supp}\,(\{\hat{y}^a\}) \subseteq \hat{\mathcal{K}}$. Note that, by the constraint condition of the P_0^i, we obtain $Y = \Phi(x(0)) + Y^a = \Phi(\hat{x}) + \hat{Y}^a$ and $\psi = \Omega x(0) = \Omega\hat{x}$, with in addition, by the objective function, we also have $|\hat{\mathcal{K}}| \leq |\mathcal{K}| \leq l$. Therefore, it is easy to see that two different initial states and two different l-sparse attack sequences explain same output and same side information, which implies that these states and attack sequences cannot be reconstructed. This contradicts the condition of Theorem 4. ∎

As indicated so far, however, P_0^i is, in general, NP-hard, and thus we relax this problem into convex one in the next subsection.

C. Relaxation to Convex Problem

Consider the following convex optimization problem for some $p \geq 1$:

$$\begin{aligned} P_{1/p}^i : \underset{\hat{x}\in\mathbb{R}^n, \hat{Y}^a \in \mathbb{R}^{m\times T}}{\text{minimize}} \quad & \left\|\hat{Y}^a\right\|_{1/p} \\ \text{subject to} \quad & Y = \Phi(\hat{x}) + \hat{Y}^a, \\ & \psi = \Omega \hat{x}. \end{aligned} \tag{31}$$

Then, the following proposition derives a necessary and sufficient condition for the equivalence of the solutions between P_0^i and $P_{1/p}^i$.

Proposition 7: Suppose that the side information ψ and Ω are given. Further, for all $z \in \mathbb{R}^n\backslash\{0\}$, suppose that (28) holds. Letting $(x(0), Y^a)$, where $x(0)$ satisfying (27) and Y^a is a matrix concatenating an l-sparse attack sequence, be the solution of P_0^i and $(\hat{x}, \hat{Y}^a)$ be the one of $P_{1/p}^i$, $(x(0), Y^a) = (\hat{x}, \hat{Y}^a)$ if and only if, for all $\mathcal{K} \subset \mathcal{S}$ with $|\mathcal{K}| = l$, the following holds for all $x \in \ker\Omega\backslash\{0\}$:

$$\sum_{i\in\mathcal{K}} \|(\Phi(x))_i\|_{\ell_p} < \sum_{i\in\mathcal{K}^\complement} \|(\Phi(x))_i\|_{\ell_p}. \tag{32}$$

Proof: For the sufficiency, we resort to a contradiction. Suppose that, for all $\mathcal{K} \subset \mathcal{S}$ with $|\mathcal{K}| = l$, (32) holds, but $(x(0), Y^a) \neq (\hat{x}, \hat{Y}^a)$, where $(x(0), Y^a)$ and $(\hat{x}, \hat{Y}^a)$ are, respectively, the solutions of P_0^i and $P_{1/p}^i$. By Proposition 6, we now know that the solution of P_0^i is the unique one, which indicates that $P_{1/p}^i$ fails to recover the unique initial state $x(0)$ satisfying (27) and l-sparse attack sequence. Thus, by the objective function of $P_{1/p}^i$, two different solutions $(x(0), Y^a)$ and $(\hat{x}, \hat{Y}^a)$ follow $\|\hat{Y}^a\|_{1/p} \leq \|Y^a\|_{1/p}$. Moreover, according to the constraint conditions of P_0^i and $P_{1/p}^i$, it follows that $Y = \Phi(x(0)) + Y^a = \Phi(\hat{x}) + \hat{Y}^a$. Here, let us define $\tilde{x} \triangleq \hat{x} - x(0) \neq 0$ following (13) and $\tilde{x} \in \ker\Omega$. Then, by the triangle inequality, we have

$$\sum_{i\in\mathcal{K}} \|(\Phi(\tilde{x}))_i\|_{\ell_p} \geq \sum_{i\in\mathcal{K}} \left(\|Y_i^a\|_{\ell_p} - \left\|\hat{Y}_i^a\right\|_{\ell_p} \right). \tag{33}$$

As given in the proof of Proposition 3, $\|Y^a\|_{1/p} = \sum_{i\in\mathcal{K}} \|Y_i^a\|_{\ell_p}$. Hence, according to the relation of $\|\hat{Y}^a\|_{1/p} \leq \|Y^a\|_{1/p}$, we obtain

$$\sum_{i\in\mathcal{K}} \|Y_i^a\|_{\ell_p} = \|Y^a\|_{1/p} \geq \left\|\hat{Y}^a\right\|_{1/p}. \tag{34}$$

Then, (33) can be formulated as

$$\begin{aligned} \sum_{i\in\mathcal{K}} \|(\Phi(\tilde{x}))_i\|_{\ell_p} &\overset{(34)}{\geq} \left\|\hat{Y}^a\right\|_{1/p} - \sum_{i\in\mathcal{K}} \left\|\hat{Y}_i\right\|_{\ell_p} \\ &= \sum_{i\in\mathcal{K}^\complement} \left\|\hat{Y}_i\right\|_{\ell_p} = \sum_{i\in\mathcal{K}^\complement} \|(\Phi(\tilde{x}))_i\|_{\ell_p}, \end{aligned}$$

where the last equality is grown out of the relation (13) and $\mathrm{supp}\,(\{y^a\}) \subseteq \mathcal{K}$. This obviously contradicts (32).

For the necessity, we again consider a contradiction, namely, assume that $(x(0), Y^a) = (\hat{x}, \hat{Y}^a)$, but (32) does not hold, where, in analogy with the previous sufficiency part, $(x(0), Y^a)$ and $(\hat{x}, \hat{Y}^a)$ are, respectively, the solutions of P_0^i and $P_{1/p}^i$, where $x(0)$ is the unique initial state satisfying (27) and Y^a is l-sparse attack matrix. In other words, there exist an index set $\mathcal{K} \subset \mathcal{S}$ with $|\mathcal{K}| = l$ and a vector $x \in \ker\Omega\backslash\{0\}$ such that

$$\sum_{i\in\mathcal{K}} \|(\Phi(x))_i\|_{\ell_p} \geq \sum_{i\in\mathcal{K}^\complement} \|(\Phi(x))_i\|_{\ell_p}. \tag{35}$$

Using a nonzero initial state $x(0) \in \ker\Omega\backslash\{0\}$, let an attack matrix be $Y^a = -(\Phi(x(0)))_i,\ i \in \mathcal{K}$. The resulting output matrix is given by $Y = \Phi(x(0)) + Y^a = (\Phi(x(0)))_i,\ i \in \mathcal{K}^\complement$. In contrast, considering another initial state $\hat{x} = 0$ and another attack matrix $\hat{Y}^a = (\Phi(x(0)))_i,\ i \in \mathcal{K}^\complement$, we have same output

matrix $Y = \Phi(\hat{x}) = \hat{Y}^a = (\Phi(x(0)))_i,\ i \in \mathcal{K}^{\complement}$. By the definitions of Y^a and $x(0)$, we have

$$\|Y^a\|_{1/p} = \sum_{i\in\mathcal{K}} \|(\Phi(x(0)))_i\|_{\ell_p} \overset{(35)}{\geq} \sum_{i\in\mathcal{K}^{\complement}} \|(\Phi(x(0)))_i\|_{\ell_p} = \left\|\hat{Y}^a\right\|_{1/p}. \quad (36)$$

This inequality indicates that $P^i_{1/p}$ fails to reconstruct the unique initial state $x(0)$ and attack matrix Y^a, i.e., $(x(0), Y^a) \neq (\hat{x}, \hat{Y}^a)$, which implies a contradiction. ■

This proposition indicates that any initial state satisfying (27) and l-sparse attack sequence can be reconstructed using a tractable convex optimization problem $P^i_{1/p}$ provided that the system satisfies (32). Note that, if we do no have any side information (i.e., $r = 0$), then we need to consider (32) for all $\mathbb{R}^n\backslash\{0\}$, which is equivalent to the previous result [19, Proposition 6] which does not consider any prior information.

Remark 4: We emphasize that the side information describes the physical attribution of the system, and this knowledge can be captured in *a priori* in many cases when we deal with some physical systems. For example, let us consider a remotely controlled unmanned aerial vehicle (UAV) whose linearized motion dynamic model is given as (1) and (2) with an external input, where its state consists of its position, velocity, attitude angle, and angular velocity [44], [45]. At initial, the UAV is landing, since, by the definition, the system is not running before $k = 0$. This indicates that the initial velocities, angular velocities, roll and pitch angles, and altitude (z-axis position) are 0. Therefore, the defender precisely knows these variables take 0 at the initial time.

VI. Numerical Simulations in Diffusion Process

A. Process Model

In this section, we illustrate the estimation performance of the proposed estimators with some prior information. This example consists of a diffusion process motivated by Remark 2, which is given as follows [46]:

$$\frac{\partial x(p,t)}{\partial t} = D\nabla^2 x(p,t),$$

where $x(p,t)$ indicates the concentration at position p at time t and D is the diffusion coefficient. Note that we assume D is independent of the position. For a further simplification, this example considers an 1-dimensional space. Regarding this process, under the Neumann boundary condition, we obtain the following time-and-position-discretized model:

$$x(k+1) = Ax(k), \quad A \triangleq I + \Delta_t A_t, \quad (37)$$

where $x(k) \triangleq [\ x(p(1), k\Delta_t), \ldots, x(p(n), k\Delta_t)\ ]^\top \in \mathbb{R}^n$ represents the concentration at sampled time k and positions $p(1), \ldots, p(n)$, which is equally spaced n positions with spacing Δ_p, Δ_t is a sampling time, and A_t is given as

$$A_t \triangleq \frac{D}{\Delta_p^2}\begin{bmatrix} -1 & 1 & & & \\ 1 & -2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & -2 & 1 \\ & & & 1 & -1 \end{bmatrix}.$$

In this example, we set $n = 25, T = 25, D = 0.7$, and $\Delta_t = 0.1$ sec. Additionally, assume that $m = 20$ and $C \in \mathbb{R}^{20\times 25}$ has iid Gaussian entries. For the readers' reference, Fig. 1 shows a behavior example of the process.

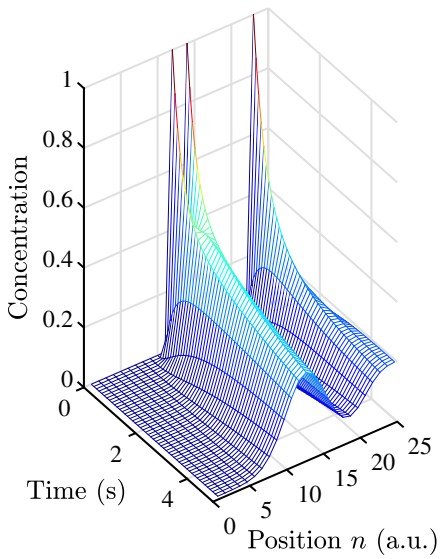

Fig. 1. Behavior example in one-dimensional diffusion process with $\bar{n}_0 = 3$.

B. Result of Sparsity Information

We first show the result that the sparsity information enhances the system resilience, that is, if the initial state is enough sparse, then the state can be reconstructed even if more sensors are compromised. To this end, for different values l and $\bar{n}_0$, we test 200 different initial conditions $x(0)$ with $\|x(0)\|_0 \leq \bar{n}_0$ and compromised sensor sets $\mathcal{K} \subseteq \mathcal{S}$ with $|\mathcal{K}| \leq l$, where $\mathcal{K}$s are chosen at random from $\mathcal{S}$. Fig. 2 illustrates the fractions of the different initial sparse conditions that are accurately reconstructed, where solid lines indicate the result of the optimization problem $P^s_{1/2}$ while dotted lines illustrate the result of $\mathcal{D}_{1/2}$, which does not consider any prior information proposed in [19]:

$$\mathcal{D}_{1/p}: \underset{\hat{x}\in\mathbb{R}^n, \hat{Y}^a\in\mathbb{R}^{m\times T}}{\text{minimize}} \quad \left\|\hat{Y}^a\right\|_{1/p} \quad \text{subject to} \quad Y = \Phi(\hat{x}) + \hat{Y}^a. \quad (38)$$

As shown in Fig. 2, if the initial state is not sparse, then the reconstruction rate of $P^s_{1/2}$ estimator deteriorates than $\mathcal{D}_{1/2}$, which follows Remark 1. On the other hand, if the initial state is sparse enough, we see that $P^s_{1/2}$ outperforms than $\mathcal{D}_{1/2}$.

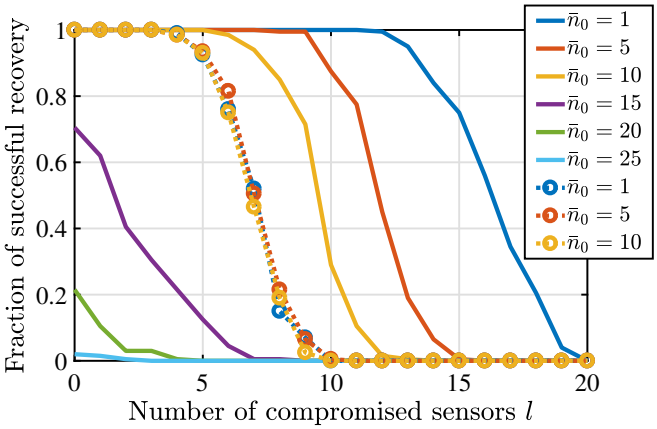

Fig. 2. Estimation performances of $P^s_{1/2}$ and $\mathcal{D}_{1/2}$ estimators in one-dimensional diffusion process with sparsity information in 200 trials. Solid lines show the result of $P^s_{1/2}$ estimator while dotted lines show the result of $\mathcal{D}_{1/2}$ estimator, which does not consider the sparsity condition.

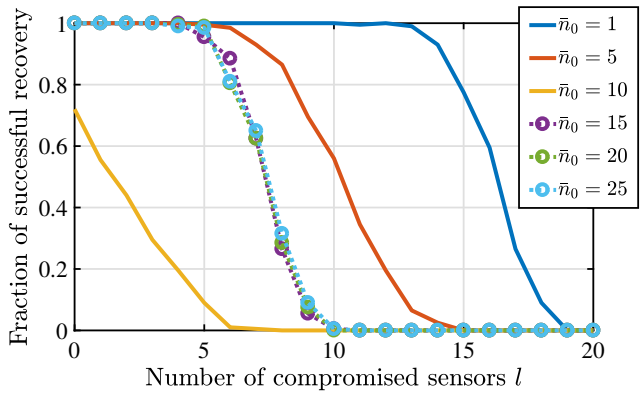

Fig. 3. Comparisons of $P^\alpha_{1/2}$ and $\mathcal{D}_{1/2}$ estimators in one-dimensional diffusion process with $(\alpha, \bar{n}_0)$-sparsity information in 200 trials. Solid lines indicate the results of $P^\alpha_{1/2}$ estimator, while dotted lines indicate the results of $\mathcal{D}_{1/2}$ estimator.

Additionally, one can confirm that the sparser the initial state is, the more accurately the state can be recovered leveraging $P^s_{1/2}$, which implies that the more initial sparsity enhances the system resilience against sensor attacks. Specifically, it is worth noticing that if the initial state is very sparse (i.e., $\bar{n}_0 = 1$), then the state reconstruction succeeds in 100% even though half of all sensors are compromised (i.e., $l = 10$).

C. Result of $(\alpha, \bar{n}_0)$-Sparsity Information

Next, we show that result of $(\alpha, \bar{n}_0)$-sparsity information. As with the previous simulation, for different values l, $\bar{n}_0$, and α, we test 200 different initial conditions $x(0)$ with $\|x(0) - \alpha \mathbf{1}_n\|_0 \leq \bar{n}_0$ and compromised sensor sets $\mathcal{K} \subseteq \mathcal{S}$ with $|\mathcal{K}| \leq l$, where αs are randomly chosen as $\alpha \in (0, 1]$. Fig. 3 shows the fractions of the different $(\alpha, \bar{n}_0)$-sparsity conditions that are correctly reconstructed, where solid lines indicate the result of the optimization problem $P^\alpha_{1/2}$ while dotted lines show the result of $\mathcal{D}_{1/2}$. This figure illustrates that if the initial state is very sparse in the sense of $(\alpha, \bar{n}_0)$-sparsity, then the reconstruction rate is enhanced. However, if the state is not sparse enough like $\bar{n}_0 = 10$, then the system resilience rather deteriorates.

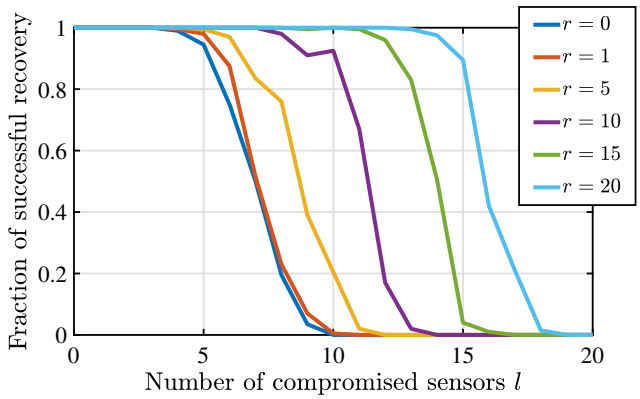

Fig. 4. Estimation performances of $P^i_{1/2}$ estimator in one-dimensional diffusion process with side information in 200 trials.

D. Result of Side Information

We finally show the resilience reinforcement result resorting to the side information. Also, for different values l and r, we test 200 different initial conditions $x(0)$ satisfying (27) and compromised sensor sets $\mathcal{K} \subseteq \mathcal{S}$ with $|\mathcal{K}| \leq l$, where r is the dimension of the side information ψ. In this example, unlike the previous simulations, we create all entries of $x(0)$ are randomly chosen from $[0, 1]$, i.e., the initial state is not sparse. Fig. 4 shows the fractions of the different side information that are correctly reconstructed using $P^i_{1/2}$, where $r = 0$ indicates no side information scenario (which is essentially equivalent to the case of $\mathcal{D}_{1/2}$). As expected, the more side information we have, the more accurate initial reconstruction achieves.

VII. Estimation Error Suppression with a priori Information

In the previous sections, we addressed that the prior information enhances the system resilience and illustrated the prior information indeed reinforces the system resilience against sensor attacks with numerical simulations. Whereas the previous sections focus on noiseless systems, this section tackles the secure estimation problem in systems with the measurement noise. In what follows, thus, let us consider the following system [23]:

$$x(k+1) = Ax(k), \tag{39}$$

$$y(k) = Cx(k) + v(k) + y^a(k), \tag{40}$$

where $v(k) \in \mathbb{R}^m$ is the measurement noise. It is assumed to be ℓ_2-bounded, and we denote the upper bound by δ, i.e., $\|v(k)\|_2 \leq \delta$, $\forall k \in \mathbb{N}_0$. This section especially devotes to derive the upper bound of the state-estimation error caused by the measurement noise. Thus, this section tackles the following problem.

Problem 2: For the system (39) and (40), provide a bound on the estimation error due to the measurement noise.

We here analyze this problem with the sparsity information of the initial state, i.e., $\bar{n}_0$, as *a priori* information. It is worth

mentioning that this analysis can be easily applied the case of $(\alpha, \bar{n}_0)$-sparse information or side information.

In analogy with (8), we have the following collected output equation for the system (39) and (40):

$$Y = \Phi(x(0)) + V + Y^a, \tag{41}$$

where

$$V \triangleq \begin{bmatrix} v(0) & \cdots & v(T-1) \end{bmatrix} \in \mathbb{R}^{m\times T}.$$

In order to estimate the initial state and attack matrix, based on (41) and the optimization problem $P^s_{1/p}$, let us consider the following optimization problem:

$$\begin{aligned} \bar{P}^s_{1/2}: \underset{\substack{\hat{x}\in\mathbb{R}^n, \hat{Y}^a\in\mathbb{R}^{m\times T}\\ \hat{V}\in\mathbb{R}^{m\times T}}}{\text{minimize}} \quad & \left\|\hat{Y}^a\right\|_{1/2} + \|\hat{x}\|_1 \\ \text{subject to} \quad & Y = \Phi(\hat{x}) + \hat{V} + \hat{Y}^a, \\ & \hat{V} \in \mathcal{V}, \end{aligned} \tag{42}$$

where $\mathcal{V}$ is the feasible region of the collected measurement noises:

$$\mathcal{V} \triangleq \left\{V\in\mathbb{R}^{m\times T}: \left\|\left(V^\top\right)_i\right\|_2 \le \delta,\ \forall i\in\{1,\dots,T\}\right\}. \tag{43}$$

For subsequent analysis, let us denote by $\tilde{x}, \tilde{Y}^a$, and $\tilde{V}$ the estimation errors as

$$\begin{aligned} &\tilde{x} \triangleq \hat{x} - x(0) \in \mathbb{R}^n, && \tilde{Y}^a \triangleq \hat{Y}^a - Y^a \in \mathbb{R}^{m\times T}, \\ &\tilde{V} \triangleq \hat{V} - V \in \mathbb{R}^{m\times T}, && \left(\hat{x}, \hat{Y}^a, \hat{V}\right) \triangleq \arg\min \bar{P}^s_{1/2}, \end{aligned} \tag{44}$$

where $x(0), Y^a$, and V are, respectively, the unique initial state, l-sparse attack matrix, and measurement noise matrix of the system. We further define the observability matrix consisting of $(A, C_\mathcal{K})$ for some $\mathcal{K}\subset\mathcal{S}$ as follows (with slight abuse of notation):

$$\mathscr{O}_\mathcal{K} \triangleq \begin{bmatrix} C_\mathcal{K}^\top & (C_\mathcal{K}A)^\top & \cdots & (C_\mathcal{K}A^{T-1})^\top \end{bmatrix}^\top \in \mathbb{R}^{|\mathcal{K}|T\times n}.$$

Note that, for any $x\in\mathbb{R}^n$ and index set $\mathcal{K}\subset\mathcal{S}$, the following inequalities hold:

$$\sum_{i\in\mathcal{K}} \|(\Phi(x))_i\|_2 \le \|\mathscr{O}_\mathcal{K} x\|_1 \le \sqrt{|\mathcal{K}|T}\,\|\mathscr{O}_\mathcal{K} x\|_2, \tag{45}$$

$$\sum_{i\in\mathcal{K}} \|(\Phi(x))_i\|_2 \ge \frac{1}{\sqrt{T}}\|\mathscr{O}_\mathcal{K} x\|_1 \ge \frac{1}{\sqrt{T}}\|\mathscr{O}_\mathcal{K} x\|_2. \tag{46}$$

A. Estimation Bound

Before continuing on, let us first obtain the following lemma regarding the estimation error of the sparse initial state.

Lemma 1: For the system (39) and (40), suppose that the initial system state satisfies $\operatorname{supp}(x(0)) \subseteq \mathcal{N}_0$ with $|\mathcal{N}_0| \le \bar{n}_0$. When sensors in a set $\mathcal{K}\subset\mathcal{S}$ are compromised, the state estimation error $\tilde{x}$ satisfies

$$\begin{aligned} \sum_{i\in\mathcal{K}} \|(\Phi(\tilde{x}))_i\|_2 + \sum_{j\in\mathcal{N}_0} |\tilde{x}_j| + 2\sigma \\ \ge \sum_{i\in\mathcal{K}^\complement} \|(\Phi(\tilde{x}))_i\|_2 + \sum_{j\in\mathcal{N}_0^\complement} |\tilde{x}_j|, \end{aligned} \tag{47}$$

where $\sigma \triangleq \max_{V\in\mathcal{V}} \|V\|_{1/2}$.

Proof: This lemma can be proved using Proposition 3 and [23, Theorem 2]. ∎

We are now ready to derive the upper bound of the state estimation error exploiting this lemma.

Theorem 5: For the system (39) and (40), suppose that the initial system state satisfies $\operatorname{supp}(x(0)) \subseteq \mathcal{N}_0$ with $|\mathcal{N}_0| \le \bar{n}_0$ and, for all $\mathcal{K}\subset\mathcal{S}$ with $|\mathcal{K}| = l$, the followings hold:

$$\mathscr{O}_{\mathcal{K}^\complement}^\top \mathscr{O}_{\mathcal{K}^\complement} - lT^2 \mathscr{O}_\mathcal{K}^\top \mathscr{O}_\mathcal{K} \ge \lambda I, \tag{48}$$

$$\frac{\lambda}{\|\mathscr{O}_{\mathcal{K}^\complement}\|_2 + \sqrt{lT}\,\|\mathscr{O}_\mathcal{K}\|_2} \ge \frac{\bar{n}_0\sqrt{T}}{\sqrt{\bar{n}_0}+1}, \tag{49}$$

for some $\lambda > 0$, where I is the $n\times n$ identity matrix. Then, when l sensors are compromised, the state estimation error $\tilde{x}$ is bounded by

$$\|\tilde{x}\|_2 \le 2\sigma\sqrt{T} \max_{|\mathcal{K}|=l} \left(\frac{\lambda}{\|\mathscr{O}_{\mathcal{K}^\complement}\|_2 + \sqrt{lT}\,\|\mathscr{O}_\mathcal{K}\|_2} - \frac{\bar{n}_0\sqrt{T}}{\sqrt{\bar{n}_0}+1}\right)^{-1}. \tag{50}$$

Proof: According to the relations (45) and (46), (47) can be rewritten as

$$\begin{aligned} &\|\mathscr{O}_{\mathcal{K}^\complement}\tilde{x}\|_2 + \sqrt{T}\left\|\tilde{x}_{\mathcal{N}_0^\complement}\right\|_1 \\ &\quad \le \sqrt{lT}\,\|\mathscr{O}_\mathcal{K}\tilde{x}\|_2 + \sqrt{T}\,\|\tilde{x}_{\mathcal{N}_0}\|_1 + 2\sigma\sqrt{T}, \\ \implies &\|\mathscr{O}_{\mathcal{K}^\complement}\tilde{x}\|_2 - \sqrt{lT}\,\|\mathscr{O}_\mathcal{K}\tilde{x}\|_2 \\ &\quad \le \sqrt{T}\,\|I_{\mathcal{N}_0}\tilde{x}\|_1 - \sqrt{T}\left\|I_{\mathcal{N}_0^\complement}\tilde{x}\right\|_1 + 2\sigma\sqrt{T}, \end{aligned} \tag{51}$$

where $I_{\mathcal{N}_0}$ and $I_{\mathcal{N}_0^\complement}$ are the matrices obtained from the $n\times n$ identity matrix by getting rid of all rows except those indexed by $\mathcal{N}_0$ and $\mathcal{N}_0^\complement$, respectively. Regarding the left-hand side of (51), it follows that

$$\begin{aligned} \|\mathscr{O}_{\mathcal{K}^\complement}\tilde{x}\|_2 - \sqrt{lT}\,\|\mathscr{O}_\mathcal{K}\tilde{x}\|_2 &= \frac{\|\mathscr{O}_{\mathcal{K}^\complement}\tilde{x}\|_2^2 - lT^2\,\|\mathscr{O}_\mathcal{K}\tilde{x}\|_2^2}{\|\mathscr{O}_{\mathcal{K}^\complement}\tilde{x}\|_2 + \sqrt{lT}\,\|\mathscr{O}_\mathcal{K}\tilde{x}\|_2} \\ &\overset{(48)}{\ge} \frac{\lambda\,\|\tilde{x}\|_2^2}{\|\mathscr{O}_{\mathcal{K}^\complement}\tilde{x}\|_2 + \sqrt{lT}\,\|\mathscr{O}_\mathcal{K}\tilde{x}\|_2} \\ &\ge \frac{\lambda\,\|\tilde{x}\|_2^2}{\left(\|\mathscr{O}_{\mathcal{K}^\complement}\|_2 + \sqrt{lT}\,\|\mathscr{O}_\mathcal{K}\|_2\right)\|\tilde{x}\|_2} \\ &= \frac{\lambda\,\|\tilde{x}\|_2}{\|\mathscr{O}_{\mathcal{K}^\complement}\|_2 + \sqrt{lT}\,\|\mathscr{O}_\mathcal{K}\|_2}. \end{aligned} \tag{52}$$

On the other hand, regarding the right-hand side of (51), by the relations between ℓ_1 and ℓ_2 norms, where

$$\frac{1}{\sqrt{n}}\|a\|_1 \le \|a\|_2 \le \|a\|_1, \quad \forall a\in\mathbb{R}^n,$$

it follows that

$$\begin{aligned}
&\sqrt{T}\left\|I_{\mathcal{N}_0}\tilde{x}\right\|_1-\sqrt{T}\left\|I_{\mathcal{N}_0^\complement}\tilde{x}\right\|_1\\
\leq&\sqrt{\bar{n}_0T}\left\|I_{\mathcal{N}_0}\tilde{x}\right\|_2-\sqrt{T}\left\|I_{\mathcal{N}_0^\complement}\tilde{x}\right\|_2\\
=&\frac{\bar{n}_0T\left\|I_{\mathcal{N}_0}\tilde{x}\right\|_2^2-T\left\|I_{\mathcal{N}_0^\complement}\tilde{x}\right\|_2^2}{\sqrt{\bar{n}_0T}\left\|I_{\mathcal{N}_0}\tilde{x}\right\|_2+\sqrt{T}\left\|I_{\mathcal{N}_0^\complement}\tilde{x}\right\|_2}\\
=&\frac{\tilde{x}^\top\left(\bar{n}_0TI_{\mathcal{N}_0}^\top I_{\mathcal{N}_0}-TI_{\mathcal{N}_0^\complement}^\top I_{\mathcal{N}_0^\complement}\right)\tilde{x}}{\sqrt{\bar{n}_0T}\left\|I_{\mathcal{N}_0}\tilde{x}\right\|_2+\sqrt{T}\left\|I_{\mathcal{N}_0^\complement}\tilde{x}\right\|_2}\\
\leq&\frac{\bar{n}_0T\left\|\tilde{x}\right\|_2^2}{\sqrt{\bar{n}_0T}\left\|I_{\mathcal{N}_0}\tilde{x}\right\|_2+\sqrt{T}\left\|I_{\mathcal{N}_0^\complement}\tilde{x}\right\|_2}\\
\leq&\frac{\bar{n}_0T\left\|\tilde{x}\right\|_2^2}{\left(\sqrt{\bar{n}_0T}+\sqrt{T}\right)\left\|\tilde{x}\right\|_2}=\frac{\bar{n}_0\sqrt{T}}{\sqrt{\bar{n}_0}+1}\left\|\tilde{x}\right\|_2. \qquad (53)
\end{aligned}$$

Thus, substituting (52) and (53) into (51), we obtain

$$\left(\frac{\lambda}{\left\|\mathscr{O}_{\mathcal{K}^\complement}\right\|_2+\sqrt{l}T\left\|\mathscr{O}_{\mathcal{K}}\right\|_2}-\frac{\bar{n}_0\sqrt{T}}{\sqrt{\bar{n}_0}+1}\right)\left\|\tilde{x}\right\|_2\leq 2\sigma\sqrt{T},$$

which derives (50). ∎

Remark 5: In the case without the prior sparsity information, the upper bound condition of (50) is given as the following (under the same condition of (48)) [23, Theorem 3]:

$$\left\|\tilde{x}\right\|_2\leq 2\sigma\sqrt{T}\max_{|\mathcal{K}|=l}\left(\frac{\lambda}{\left\|\mathscr{O}_{\mathcal{K}^\complement}\right\|_2+\sqrt{l}T\left\|\mathscr{O}_{\mathcal{K}}\right\|}\right)^{-1}$$

Thus, one can easily observe that, taking the sparsity information into account, the term of $N\triangleq\frac{\bar{n}_0\sqrt{T}}{\sqrt{\bar{n}_0}+1}$ is involved in the upper bound condition. The sparser the initial state is (or the smaller $\bar{n}_0$ is), the smaller N is, which implies the sparsity information turns out the decreasing of the upper bound of $\|\tilde{x}\|_2$. Therefore, the sparser initial state suppresses the estimation error due to the measurement noises as well.

B. Numerical Simulations

This subsection illustrates that the prior information suppresses the effect of measurement noise. We here also resort to the diffusion process developed in Section VI. Similarly, we set $n=25, m=20, T=25, D=0.7$, and $\Delta_t=0.1$ sec. Also, assume that $C\in\mathbb{R}^{20\times 25}$ has iid Gaussian entries and the noise bound satisfies $\delta=1$. For different initial sparsity conditions $\bar{n}_0=1,5,10$, we also test 200 scenarios with $l=8$ sensor attacks and randomly chosen compromised sensor sets $\mathcal{K}\subset\mathcal{S}$ with $|\mathcal{K}|\leq 8$.

Figs. 5 and 6 show the histograms of estimation errors in different sparsity conditions, where Fig. 5 shows the results of the optimization problem $\bar{P}^s_{1/2}$ while Fig. 6 illustrates the

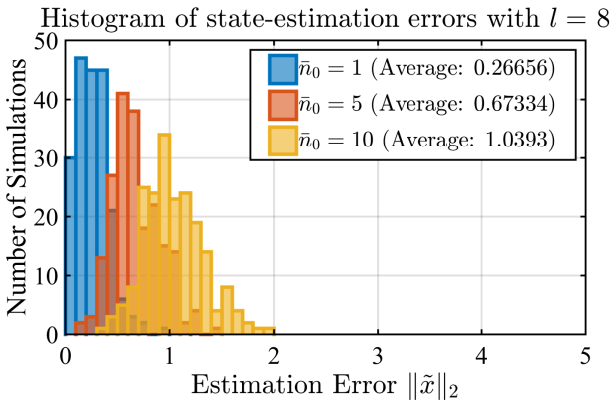

Fig. 5. Estimation performance of $\bar{P}^s_{1/2}$ estimator in one-dimensional diffusion process with different sparsity information in 200 trials.

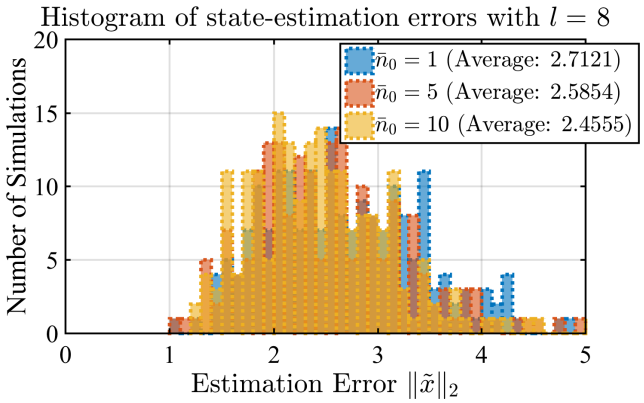

Fig. 6. Estimation performance of $\bar{\mathcal{D}}_{1/2}$ estimator in one-dimensional diffusion process with different sparsity information in 200 trials.

ones of $\bar{\mathcal{D}}_{1/2}$, which does not consider any prior information proposed in [23]:

$$\begin{aligned}
\bar{\mathcal{D}}_{1/2}:\ \underset{\substack{\hat{x}\in\mathbb{R}^n,\hat{Y}^a\in\mathbb{R}^{m\times T}\\ \hat{V}\in\mathbb{R}^{m\times T}}}{\text{minimize}}\ &\left\|\hat{Y}^a\right\|_{1/2}\\
\text{subject to}\ \ &Y=\Phi\left(\hat{x}\right)+\hat{V}+\hat{Y}^a,\\
&\hat{V}\in\mathcal{V}. \qquad (54)
\end{aligned}$$

As illustrated in Fig. 5, we observe that the sparsity information reduces the estimation error caused by the measurement noise. Further, comparing Fig. 5 with Fig. 6, $\bar{P}^s_{1/2}$ obviously outperforms $\bar{\mathcal{D}}_{1/2}$ in terms of the estimation error. This indicates that, by taking the sparsity information into account explicitly, the state-estimation error is suppressed by the information. Therefore, we confirm that the sparsity information also reinforces the system resilience in the presence of the measurement noise.

VIII. Conclusion

In this paper, we have shown that the prior information of the estimated state enhances the system resilience, that is, even if more sensors are compromised, when the information of the state is given, then one can uniquely recover the state and attack. We especially have focused on the sparsity information, $(\alpha,\bar{n}_0)$-sparsity information, and side

information. In each scenario, we derived a necessary and sufficient condition for the state reconstruction, and then, an estimator with the information was developed. Under a certain condition, furthermore, it was presented that the solution of each estimator coincides with the one of a tractable convex estimator. Indeed, all conditions for the state reconstruction have indicated even if more sensors are compromised, the state can be recovered resorting to the prior information, which indicates that the system resilience against sensor attacks is reinforced by the information. Additionally, we extend this analysis to noisy systems, and we have shown that the prior information suppresses the state-estimation error caused by the bounded measurement noise. We finally illustrated the resilient reinforcement and error-suppression results using a diffusion process model.

REFERENCES

[1] K. D. Kim and P. R. Kumar, "Cyber-physical systems: A perspective at the centennial," *Proc. IEEE*, vol. 100, no. Special Centennial Issue, pp. 1287–1308, 2012.

[2] J. Sztipanovits, X. Koutsoukos, G. Karsai, N. Kottenstette, P. Antsaklis, V. Gupta, B. Goodwine, J. Baras, and S. Wang, "Toward a science of cyber-physical system integration," *Proc. IEEE*, vol. 100, no. 1, pp. 29–44, 2012.

[3] F. Pasqualetti, F, Dörfler, and F. Bullo, "Control-theoretic methods for cyberphysical security: Geometric principles for optimal cross-layer resilient control systems," *IEEE Control Syst. Mag.*, vol. 35, no. 1, pp. 110–127, 2015.

[4] S. Sridhar, A. Hahn, and M. Govindarasu, "Cyber-physical sysytem security for the electric power grid," *Proc. IEEE*, vol. 100, no. 1, pp. 210–224, 2012.

[5] A. Cardenas, S. Amin, B. Sinopoli, A. Giani, A. Perrig, and S. Sastry, "Challenges for securing cyber physical systems," in *Proc. Workshop Future Directions Cyber-Phys. Syst. Security*, Newark, NJ, 2009.

[6] R. Langner, "Stuxnet: Dissecting a cyberwarfare weapon," *IEEE Secur. Privacy*, vol. 9, no. 3, pp. 49–51, 2011.

[7] J. Slay and M. Miller, "Lesson learned from the Maroochy water branch," *Critical Infrastructure Protection*, vol. 253, pp. 73–82, 2007.

[8] T. Pultarova, "Cyber security: Ukraine grid hack is wake-up call for network operators [News Briefing]," *Eng. & Technology*, vol. 11, no. 1, pp. 12–13, 2016.

[9] Y. Mo, R. Chabukswar, and B. Sinopoli, "Detecting integrity attacks on SCADA systems," *IEEE Trans. Control Syst. Technol.*, vol. 22, no. 4, pp. 1396–1407, 2014.

[10] R. H. Smith, "Covert misappropriation of networked control systems: Presenting a feedback structure," *IEEE Control Syst. Mag.*, vol. 35, no. 1, pp. 82–92, 2015.

[11] A. Teixeira, I. Shames, H. Sandberg, and K. H. Johansson, "A secure control framework for resource-limited adversaries," *Automatica*, vol. 51, pp. 135–148, 2015.

[12] F. Pasqualetti, F, Dörfler, and F. Bullo, "Attack detection and identification in cyber-physical systems," *IEEE Trans. Autom. Control*, vol. 58, no. 11, pp. 2715–2729, 2013.

[13] Y. Mo and B. Sinopoli, "On the performance degradation of cyber-physical systems under stealthy integrity attacks," *IEEE Trans. Autom. Control*, vol. 61, no. 9, pp. 2618–2624, 2016.

[14] J. Back, J. Kim, C. Lee, G. Park, and H. Shim, "Enhancement of security against zero dynamics attack via generalized hold," in *Proc. 56th IEEE Conf. Decision and Control*, Melbourne, Australia, 2017, pp. 1350–1355.

[15] T. Shinohara and T. Namerikawa, "Manipulative zero-stealthy attacks in cyber-physical systems: Existence space of feasible attack objectives," in *Proc. 1st IEEE Conf. Control Technology and Applicat.*, Kohala Coast, HI, 2017, pp. 1123–1128.

[16] T. Shinohara and T. Namerikawa, "On the vulnerabilities due to manipulative zero-stealthy attacks in cyber-physical systems," *SICE J. Control, Measurement, and Syst. Integration*, vol. 10, no. 6, pp. 563–570, 2017.

[17] S. Weerakkody, X. Liu, S. H. Son, and B. Sinopoli, "A graph-theoretic characterization of perfect attackability for secure design of distributed control systems," *IEEE Trans. Control Netw. Syst.*, vol. 4, no. 1, pp. 60–70, 2017.

[18] F. Miao, Q. Zhu, M. Pajic, and G. J. Pappas, "Coding schemes for securing cyber-physical systems against data injection attacks," *IEEE Trans. Control Netw. Syst.*, vol. 4, no. 1, pp. 106–117, 2017.

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

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

[21] M. S. Chong, M. Wakaiki, and J. P. Hespanha, "Observability of linear systems under adversarial attacks," in *Proc. 2015 Amer. Control Conf.*, Chicago, IL, 2015, pp. 2439–2444.

[22] M. Pajic, P. Tabuada, I. Lee, and G. J. Pappas, "Attack-resilient state estimation in the presence of noise," in *Proc. IEEE 54th Conf. Decision and Control*, Osaka, Japan, 2015, pp. 5827-5832.

[23] M. Pajic, I. Lee, and G. J. Pappas. "Attack-resilient state estimation for noisy dynamical systems," *IEEE Trans. Control Netw. Syst*, vol. 4, no. 1, pp. 82–92, 2017.

[24] M. Pajic, J. Weimer, N. Bezzo, O. Sokolsky, G. J. Pappas, and I. Lee, "Design and implementation of attack-resilient cyber-physical systems," *IEEE Control Syst. Mag.*, vol. 37, no. 2, pp. 66–81, 2017.

[25] C. Lee, H. Shim, and Y. Eun, "Secure and robust state estimation under sensor attacks, measurement noises, and process disturbances: Observer-based combinational approach," in *Proc. 2015 European Control Conf.*, Linz, Austria, 2015, pp. 1872–1875.

[26] Y. Nakahira and Y. Mo, "Dynamic state estimation in the presence of compromised sensory data," in *Proc. 54th IEEE Conf. Decision and Control*, Osaka, Japan, 2015, pp. 5808–5813.

[27] T. Shinohara and T. Namerikawa, "Reach set-based attack resilient state estimation against omniscient adversaries," in *Proc. 2018 Amer. Control Conf.*, Milwaukee, WI, 2018, pp. 5813–5818.

[28] Y. Shoukry, P. Nuzzo, A. Puggelli, A. L. Sangiovanni-Vincentelli, S. A. Seshia, and P. Tabuada, "Secure estimation for cyber physical systems under sensor attacks: A satisfiability modulo theory approach," *IEEE Trans. Autom. Control*, vol. 62, no. 10, pp. 4917–4932, 2017.

[29] 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," in *Proc. 56th IEEE Conf. Decision and Control*, Melbourne, Australia, 2017, pp. 157–162.

[30] Y. Shoukry, P. Nuzzo, A. Puggelli, 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," in *Proc. 54th IEEE Conf. Decision and Control*, Osaka, Japan, 2015, pp. 3804–3809.

[31] 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," in *Proc. 55th IEEE Conf. Decision and Control*, Las Vegas, NV, 2016, pp. 1297–1302.

[32] 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 Trans. Control Netw. Syst.*, 2018. (early access)

[33] H. L. Trentleman, A. A. Stoorvogel, and M. Hautus, *Control Theory for Linear Systems*. Springer, 2001.

[34] E. J. Candés and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[35] D. L. Donoho, M. Elad, and V. N. Temlyakov "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 6–18, 2006.

[36] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ_1 minimization," *Proc. Nat. Academy of Sci. of USA*, vol. 100, no. 5, pp. 2197–2202, 2003.

[37] I. F. Gorodnitsky and B. D. Rao, "Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 600–616. 1997.

[38] E. Elhamifar and R. Vidal, "Block-sparse recovery via convex optimization," *IEEE Trans. Signal Process.*, vol. 60, no. 8, pp. 4094–4107, 2012.

[39] B. K. Natarajan, "Sparse approximate solutions to linear systems," *SIAM J. Comput.*, vol. 24, no. 2, pp. 227–234, 1995.

[40] A. M. Tillmann and M. E. Pfetsch, "The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 1248–1259, 2014.

[41] P. Wang, M. C. Conzález, C. A. Hidalgo, and A.-L. Barabási, "Understanding the spreading patterns of mobile phone viruses," *Science*, vol. 324, no. 5930, pp. 1071–1076, 2009.

[42] Z.-L. Hu, Z. Shen, C.-B. Tang, B.-B. Xie, and J.-F. Lu, "Localization of diffusion sources in complex networks with sparse observations," *Physics Lett. A*, vol. 382, no. 14, pp. 931–937, 2018.
[43] Y. Chen, S. Kar, and J. M. F. Moura, "Dynamic attack detection in cyber-physical systems with side initial state information," *IEEE Trans. on Autom. Control*, vol. 62, no. 6, pp. 4618–4624, 2017.
[44] Z. T. Dyrek, A. M. Annaswamy, and E. Lavretsky, "Araptive control of quadrotor UAVs: A design trade study with flight evaluations," *IEEE Trans. Control Syst. Technol.*, vol. 21, no. 4, pp. 1400–1406, 2013.
[45] T. Namerikawa, Y. Kuriki, and A. Khalifa, "Consensus-based cooperative formation control for multiquadcopter system with unidirectional network connections," *J. Dyn. Sys., Meas., Control*, vol. 140, no. 4, pp. 044502-1–8, 2018.
[46] J. Crank, *The Mathematics of Diffusion*. Oxford University Press, 1975.

Takumi Shinohara received the B.E. and M.E. degrees from Keio University, Tokyo, Japan, in 2016 and 2018, respectively. His research interests include system security and secure state estimation.

Toru Namerikawa (M'94–SM'19) received the B.E., M.E., and Ph.D. in electrical and computer engineering from Kanazawa University, Japan, in 1991, 1993 and 1997, respectively. From 1994 until 2002, he was with Kanazawa University as an Assistant Professor. From 2002 until 2005, he was with the Nagaoka University of Technology as an Associate Professor, Niigata, Japan. From 2006 until 2009, he was with Kanazawa University again. In April 2009, he joined Keio University, Yokohama, Japan, where he is currently a Professor at Department of System Design Engineering, Keio University. He held visiting positions at Swiss Federal Institute of Technology in Zurich in 1998, University of California, Santa Barbara in 2001, University of Stuttgart in 2008 and Lund University in 2010. He received 2014 Pioneer Technology Award from SICE Control Division and 2017 Outstanding Paper Award from SICE. His main research interests are robust control, distributed and cooperative control and their application to power network systems.

Zhihua Qu (M'90–SM'93–F'09) received the Ph.D. degree in electrical engineering from the Georgia Institute of Technology, Atlanta, in June 1990. Since then, he has been with the University of Central Florida (UCF), Orlando. Currently, he is the SAIC Endowed Professor in College of Engineering and Computer Science, a Pegasus Professor and the Chair of Electrical and Computer Engineering, and the Director of FEEDER Center. His areas of expertise are nonlinear systems and control, with applications to autonomous systems and energy/power systems. His recent work focuses upon control, optimization, and plug-&-play resilient operation of networked systems.