On Redundant Observability: From Security Index to Attack Detection and Resilient State Estimation

Chanhwa Lee, Hyungbo Shim, *Senior Member, IEEE*, and Yongsoon Eun, *Member, IEEE*

***Abstract*—The security of control systems under sensor attacks is investigated. Redundant observability is introduced, explaining existing security notions including the security index, attack detectability, and observability under attacks. Equivalent conditions between redundant observability and existing notions are presented. Based on a bank of partial observers utilizing Kalman decomposition and a decoder exploiting redundancy, an estimator design algorithm is proposed enhancing the resilience of control systems. This scheme substantially improves computational efficiency utilizing far less memory.**

***Index Terms*—Analytical redundancy, attack detection, attack resilience, cyber-physical systems, resilient state estimation, security index.**

Notation: The subset of natural numbers, $\{1, 2, \ldots, \mathsf{p}\} \subset \mathbb{N}$, is denoted by $[\mathsf{p}]$. The cardinality of a set S is denoted by $|S|$ and the support of a vector $y \in \mathbb{C}^{\mathsf{p}}$ is defined as $\mathsf{supp}(y) := \{\mathsf{i} \in [\mathsf{p}] : y_{\mathsf{i}} \neq 0\}$ where y_{i} is the i-th element of y. The cardinality of $\mathsf{supp}(y)$ defines the ℓ_0 norm of a vector y, i.e., $\|y\|_0 := |\mathsf{supp}(y)|$. A vector y is said to be q-sparse if $\|y\|_0 \leq \mathsf{q}$. The set $\Sigma_{\mathsf{q}} := \{y \in \mathbb{C}^{\mathsf{p}} : \|y\|_0 \leq \mathsf{q}\}$ denotes the set of all q-sparse vectors. The 2-norm of a vector y is defined as $\|y\|_2 := \sqrt{y^* y}$, where y^* is the Hermitian of y.

Assume that a vector $y \in \mathbb{C}^{\mathsf{p}}$ and a subset $\Lambda \subset [\mathsf{p}]$ of indices are given. We use the notation $y_\Lambda \in \mathbb{C}^{\mathsf{p}}$ to denote that y_Λ is obtained by setting the elements of y indexed by $\Lambda^c := [\mathsf{p}] \backslash \Lambda = \{\mathsf{i} \in [\mathsf{p}] : \mathsf{i} \notin \Lambda\}$ to zero. Similar notation is used for a matrix $C \in \mathbb{R}^{\mathsf{p} \times \mathsf{n}}$. The matrix obtained by setting the rows of C indexed by Λ^c to zero, is denoted as $C_\Lambda \in \mathbb{R}^{\mathsf{p} \times \mathsf{n}}$. Sometimes the notation will be slightly modified to $y_\Lambda^\pi \in \mathbb{C}^{|\Lambda|}$ (or $C_\Lambda^\pi \in \mathbb{R}^{|\Lambda| \times \mathsf{n}}$), which denotes the vector y (or the matrix C) whose elements (or rows) not corresponding to the index set Λ are actually eliminated.

For a given index $\mathsf{i} \in [\mathsf{p}]$, the index set $\Gamma_{\mathsf{i}}^{\mathsf{n}} \subset [\mathsf{np}]$ represents $\{\mathsf{n}(\mathsf{i}-1)+1, \mathsf{n}(\mathsf{i}-1)+2, \ldots, \mathsf{ni}\}$. Similarly, for a given index set $\Lambda \subset [\mathsf{p}]$, the index set $\Lambda^{\mathsf{n}} \subset [\mathsf{np}]$ denotes $\bigcup_{\mathsf{i} \in \Lambda} \Gamma_{\mathsf{i}}^{\mathsf{n}}$. A vector $z \in \mathbb{C}^{\mathsf{np}}$ of length np can be split into p column vectors of length n, i.e., $z = [z_1^{\mathsf{n}\top}\ z_2^{\mathsf{n}\top}\ \cdots\ z_{\mathsf{p}}^{\mathsf{n}\top}]^\top \in \mathbb{C}^{\mathsf{np}}$, where $z_{\mathsf{i}}^{\mathsf{n}} \in \mathbb{C}^{\mathsf{n}}$ represents the i-th split column vector of length n in z. Then we call z an n-stacked vector. With the index set $\Gamma_{\mathsf{i}}^{\mathsf{n}}$ defined above, it follows that $z_{\mathsf{i}}^{\mathsf{n}} = z_{\Gamma_{\mathsf{i}}^{\mathsf{n}}}^{\pi} \in \mathbb{C}^{\mathsf{n}}$. The ($\mathsf{n}$-stacked) support of $z \in \mathbb{C}^{\mathsf{np}}$ is defined as $\mathsf{supp}^{\mathsf{n}}(z) := \{\mathsf{i} \in [\mathsf{p}] : z_{\mathsf{i}}^{\mathsf{n}} \neq 0_{\mathsf{n}\times 1}\}$ and its cardinality defines the (n-stacked) ℓ_0 norm of z, i.e., $\|z\|_{0^{\mathsf{n}}} := |\mathsf{supp}^{\mathsf{n}}(z)|$. Similarly to the usual vector case, an n-stacked vector z is said to be (n-stacked) q-sparse when it holds that $\|z\|_{0^{\mathsf{n}}} \leq \mathsf{q}$, and the set $\Sigma_{\mathsf{q}}^{\mathsf{n}} := \{z \in \mathbb{C}^{\mathsf{np}} : \|z\|_{0^{\mathsf{n}}} \leq \mathsf{q}\}$ denotes the set of all (n-stacked) q-sparse vectors.

For a matrix $C \in \mathbb{R}^{\mathsf{p} \times \mathsf{n}}$, the cospark of C is defined as $\mathsf{cospark}(C) := \min_{x \in \mathbb{R}^{\mathsf{n}},\, x \neq 0_{\mathsf{n}\times 1}} \|Cx\|_0$ and the (n-stacked) cospark of a matrix $\Phi \in \mathbb{R}^{\mathsf{np} \times \mathsf{n}}$ is similarly defined as $\mathsf{cospark}^{\mathsf{n}}(\Phi) := \min_{x \in \mathbb{R}^{\mathsf{n}},\, x \neq 0_{\mathsf{n}\times 1}} \|\Phi x\|_{0^{\mathsf{n}}}$. Subspaces $\mathcal{R}(C)$ and $\mathcal{N}(C)$ denote the range space and the null space of C, respectively. The induced matrix 2-norm of a matrix C is defined as $\|C\|_2 := \sqrt{\lambda_{\max}(C^\top C)} = \sigma_{\max}(C)$, where $\lambda_{\max}(\cdot)$ and $\sigma_{\max}(\cdot)$ denote the maximum eigenvalue and the maximum singular value, respectively. In addition, $\sigma_{\min}(\cdot)$ is used to denote the minimum singular value and $C^\dagger$ is the pseudoinverse of C. Finally, the set of normalized eigenvectors of a square matrix $A \in \mathbb{R}^{\mathsf{n} \times \mathsf{n}}$ is denoted as $\mathcal{V}(A) := \{v \in \mathbb{C}^{\mathsf{n}} : Av = \lambda v \text{ for some } \lambda \in \mathbb{C},\ \|v\|_2 = 1\}$.

I. Introduction

The reliability of systems in various circumstances is one of the main concerns for control engineers, and thus robust and fault-tolerant control methods have been developed to cope with model uncertainties, external disturbances, and failures in system components. Recently, new threats or vulnerabilities caused by malicious attacks have been reported as advances in computers and communications increase the connectivity and openness of systems [2]. Therefore, the resilience of control systems to attack has become a critical system design consideration [3]–[5] and the security problems of the system whose measurements are compromised by adversaries have been studied actively because sensors are one of the most vulnerable points for the security of control systems [6]–[15].

In this paper, we consider a discrete-time linear time invariant (LTI) system under sensor attacks written as

$$\mathcal{P} : \begin{cases} x(k+1) = Ax(k) + Bu(k) + d(k) \\ \bar{y}(k) = y(k) + a(k) = Cx(k) + n(k) + a(k) \end{cases} \tag{1}$$

where $x \in \mathbb{R}^{\mathsf{n}}$ denotes the state variables, $u \in \mathbb{R}^{\mathsf{m}}$ denotes the control inputs, $y \in \mathbb{R}^{\mathsf{p}}$ denotes the attack-free sensor outputs, and $\bar{y} \in \mathbb{R}^{\mathsf{p}}$ denotes the measurement data under attack signals. The dynamics are disrupted by the process disturbance $d \in \mathbb{R}^{\mathsf{n}}$ and sensors are corrupted by the sensor attack $a \in \mathbb{R}^{\mathsf{p}}$ as well as the measurement noise $n \in \mathbb{R}^{\mathsf{p}}$. There is a total of p sensors that measure the system outputs and the i-th measurement data at time k is denoted by $\bar{y}_{\mathsf{i}}(k) = c_{\mathsf{i}} x(k) + n_{\mathsf{i}}(k) + a_{\mathsf{i}}(k)$, where c_{i} is the i-th row of C. It is

Manuscript received November 13, 2017; revised February 6, 2018; accepted May 3, 2018. Date of publication May 16, 2018; date of current version January 28, 2019. This work was supported in part by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (2014-0-00065, Resilient Cyber-Physical Systems Research) and in part by Global Research Laboratory Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (NRF-2013K1A1A2A02078326). This paper was presented in part at the 14th European Control Conference, Linz, Austria, Jul. 2015 as [1]. Recommended by Associate Editor Z. Gao. *(Corresponding author: Hyungbo Shim.)*

C. Lee is with Research & Development Division, Hyundai Motor Company, South Korea (e-mail: chanhwa.lee@gmail.com).

H. Shim is with ASRI, Department of Electrical and Computer Engineering, Seoul National University, South Korea (e-mail: hshim@snu.ac.kr).

Y. Eun is with Department of Information & Communication Engineering, DGIST, South Korea (e-mail: yeun@dgist.ac.kr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2018.2837107

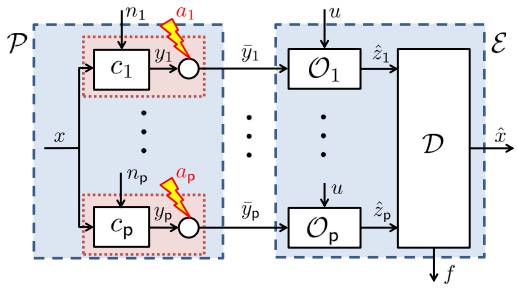

Fig. 1. Configuration of the plant $\mathcal{P}$ and the state estimator $\mathcal{E}$.

assumed that the disturbances/noises are uniformly bounded, and the attacks can compromise up to q out of p sensor outputs, as follows.

Assumption 1: The process disturbance d and each measurement noise n_i are uniformly bounded, i.e.,

$$\|d(k)\|_2 \le d_{\max},\ \|n_i(k)\|_2 \le n_{\max},\ {}^\forall k \ge 0,\ {}^\forall i \in [\mathsf{p}]. \qquad \Diamond$$

Assumption 2: There exist at least $\mathsf{p}-\mathsf{q}$ sensors that are not attacked for all $k \ge 0$, i.e.,

$$\left|\{i \in [\mathsf{p}] : a_i(k) = 0,\ {}^\forall k \ge 0\}\right| \ge \mathsf{p}-\mathsf{q}. \qquad \Diamond$$

The primary objective of this paper is to design an estimator $\mathcal{E}$ that detects the attacked sensors and estimates the state $x(k)$ of the given system $\mathcal{P}$ under Assumptions 1 and 2. To this end, we first characterize the conditions under which the attack can be detected and the state of $\mathcal{P}$ can be estimated correctly. Then, we construct an attack-resilient estimator $\mathcal{E}$ that is composed of p partial observers[1] $\mathcal{O}_i$ and a decoder $\mathcal{D}$ as shown in Fig. 1. In other words, the characterization of observability with unknown signal a and construction of the estimator are the two main topics of this paper.[2]

In the first part of this paper, a vulnerability analysis is conducted. Fundamental limitations such as attack detectability (and identifiability) conditions have been investigated in [4] and the attack detectability is quantified by the security index [7], which is the minimum number of attacks to remain undetectable. This security index concept for a static output map is generalized to a dynamical system under sensor attacks in [8]. We have carefully explained the relationship between these fundamental limitations and the *redundant observability*, which is a kind of the analytical redundancy in measurements and will be formally defined in Section III. Furthermore, equivalent conditions between them are also presented.

In the second part of this paper, we propose a resilient and robust state estimation scheme. Compared with the existing resilient estimation algorithms in [9]–[15], the advantages of our scheme are as follows. First, it does not require any additional restrictive conditions other than the redundant observability (compared with [9], [11], [13], [15]). Second, an observer-based algorithm makes it possible to estimate the current state, not the initial state or delay information (compared with [9], [10], [12]). Third, the scheme is robust in the sense that a bound on estimation error is explicitly derived from system parameters (compared with [9]–[14]). Finally, the scheme requires less computational effort and less memory owing to the reduction in time and space complexity (compared with [14]).

The rest of the paper is organized as follows. Section II presents the theoretical background of static error correcting problems for a stacked vector case. We then present the relationship between redundant observability and security-related concepts such as dynamic security index, attack detectability, and observability under attacks in Section III. In addition, partial observers using the Kalman observability decomposition are designed and the overall resilient and robust estimation scheme is presented in Section IV. Finally, simulation results with a three inertia system are given in Section V and we provide concluding remarks in Section VI.

II. Static Error Correction for Stacked Vector

In this section, a static error correcting algorithm is studied that will play a key role for constructing the decoder $\mathcal{D}$ in the estimator $\mathcal{E}$. In particular, we solve a particular problem: Given a matrix $\Phi \in \mathbb{R}^{\mathsf{np}\times\mathsf{n}}$, recover an unknown vector $x \in \mathbb{R}^{\mathsf{n}}$ from the known measurement $\hat{z}$ given by[3]

$$\hat{z} = \Phi x + v + e \in \mathbb{R}^{\mathsf{np}} \tag{2}$$

where the n-stacked vector $\hat{z} \in \mathbb{R}^{\mathsf{np}}$ is corrupted by two more unknown vectors $v \in \mathbb{R}^{\mathsf{np}}$ and $e \in \mathbb{R}^{\mathsf{np}}$. The vector v represents noise and is assumed to have bounded magnitude. The vector e is called error, and it corresponds to an attack signal whose magnitude can be arbitrarily large but is assumed to be sparse. The matrix Φ is called a *coding matrix*.

A. Error Detectability and Detection Scheme

One should be able to detect the existence of an error to reconstruct the original state vector x. Thus, we start this subsection by introducing the notion of *error detectability* when the measurement $\hat{z}$ in (2) is noise free (i.e., $v = 0_{\mathsf{np}\times 1}$).

Definition 1: A coding matrix $\Phi \in \mathbb{R}^{\mathsf{np}\times\mathsf{n}}$ is said to be *(n-stacked) q-error detectable* if, for all $x, x' \in \mathbb{R}^{\mathsf{n}}$ and $e \in \Sigma^{\mathsf{n}}_{\mathsf{q}}$ such that $\Phi x + e = \Phi x'$, it holds that $x = x'$.

Therefore, the matrix $\Phi \in \mathbb{R}^{\mathsf{np}\times\mathsf{n}}$ is not (n-stacked) q-error detectable if and only if there are two different x and x' in $\mathbb{R}^{\mathsf{n}}$, and e in $\Sigma^{\mathsf{n}}_{\mathsf{q}}$ such that $\Phi x + e = \Phi x'$. Now, two more equivalent conditions that characterize the error detectability of a coding matrix Φ are given.

Proposition 1: The following are equivalent:

1) The matrix $\Phi \in \mathbb{R}^{\mathsf{np}\times\mathsf{n}}$ is (n-stacked) q-error detectable.
2) For every set $\Lambda \subset [\mathsf{p}]$ satisfying $|\Lambda| \ge \mathsf{p}-\mathsf{q}$, $\Phi_{\Lambda^{\mathsf{n}}}$ (or, equivalently, $\Phi^{\pi}_{\Lambda^{\mathsf{n}}}$) has full column rank.
3) For any $x \in \mathbb{R}^{\mathsf{n}}$ where $x \ne 0_{\mathsf{n}\times 1}$, $\|\Phi x\|_{0^{\mathsf{n}}} > \mathsf{q}$.

Proof: 1) $\Rightarrow$ 2): Suppose that 2) does not hold, i.e., there exists an index set $\Lambda \subset [\mathsf{p}]$ with $|\Lambda| \ge \mathsf{p}-\mathsf{q}$ and $x \ne 0_{\mathsf{n}\times 1}$ such that $\Phi_{\Lambda^{\mathsf{n}}} x = 0_{\mathsf{np}\times 1}$. Then it follows that $\|e\|_{0^{\mathsf{n}}} \le \mathsf{q}$, where $e := -\Phi x$. Thus, $\Phi x + e = \Phi 0_{\mathsf{n}\times 1}$ and Φ is not q-error detectable.

2) $\Rightarrow$ 3): Suppose, for the sake of contradiction, that there exists $x \ne 0_{\mathsf{n}\times 1}$ such that $\|\Phi x\|_{0^{\mathsf{n}}} \le \mathsf{q}$. Let Λ be the complement of $\mathsf{supp}^{\mathsf{n}}(\Phi x)$, i.e., $\Lambda = (\mathsf{supp}^{\mathsf{n}}(\Phi x))^c$. Then it is obvious that $|\Lambda| \ge \mathsf{p}-\mathsf{q}$ and $\Phi_{\Lambda^{\mathsf{n}}} x = 0_{\mathsf{np}\times 1}$. This contradicts the full column rank condition of $\Phi_{\Lambda^{\mathsf{n}}}$ in 2).

3) $\Rightarrow$ 1): We again prove it by contradiction. Suppose that Φ is not q-error detectable. That is, there exist $x, x' \in \mathbb{R}^{\mathsf{n}}$ satisfying $x \ne x'$,

[1] In this paper, the terms "observer" and "estimator" are used to indicate the block of $\mathcal{O}_i$ and $\mathcal{E}$ in Fig. 1, respectively. That is, two terminologies should be distinguished.

[2] From the control theoretic perspective, strong observability [16] and unknown input observer (UIO)-based fault estimation [17] may be closely related to the subject of interest here. If we consider the output equation $\bar{y}(k) = Cx(k) + n(k) + I_\Lambda a(k)$ (instead of imposing Assumption 2 on a in (1)) where $I \in \mathbb{R}^{\mathsf{p}\times\mathsf{p}}$ is an identity matrix and $\Lambda \subset [\mathsf{p}]$ is any index set satisfying $|\Lambda| \le \mathsf{q}$, then, as mentioned in [11], the problem of interest is strong observability for any q-sparse identity matrix I_Λ, and the design of a UIO-based estimator for unknown Λ.

[3] Later on, the analysis in Section III is performed based on the measurement (7) and the design in Section IV is carried out based on the estimation error (19). Note that both equations are in the form of (2).

and $e \in \Sigma_{\mathsf{q}}^{\mathsf{n}}$ such that $\Phi x + e = \Phi x'$. It follows from $x' - x \neq 0_{\mathsf{n}\times 1}$ and $e \in \Sigma_{\mathsf{q}}^{\mathsf{n}}$ that $\|\Phi(x'-x)\|_{0^{\mathsf{n}}} = \|e\|_{0^{\mathsf{n}}} \leq \mathsf{q}$. Thus, condition 3) does not hold. ■

Remark 1: In Proposition 1, condition 2) relates q-error detectability to the left invertibility of Φ. That is, Φ remains left invertible even if any (n-stacked) q row blocks are eliminated. We may call this property *q-redundant left invertibility*. On the other hand, condition 3) establishes the link between the error detectability and the cospark of a coding matrix. More specifically, Φ is q-error detectable if and only if its cospark is larger than q, i.e., $\mathsf{cospark}^{\mathsf{n}}(\Phi) > \mathsf{q}$. ♦

The equivalence conditions in Proposition 1 lead to a criterion of q-sparse error detection based on a residual signal

$$r := \hat{z} - \Phi\Phi^{\dagger}\hat{z} = \left(I_{\mathsf{np}\times\mathsf{np}} - \Phi(\Phi^{\top}\Phi)^{-1}\Phi^{\top}\right)\hat{z}. \tag{3}$$

Lemma 1: For the measurement $\hat{z} = \Phi x + e$ where $\Phi \in \mathbb{R}^{\mathsf{np}\times\mathsf{n}}$ is (n-stacked) q-error detectable, $x \in \mathbb{R}^{\mathsf{n}}$, and $e \in \Sigma_{\mathsf{q}}^{\mathsf{n}}$, let $r = \hat{z} - \Phi\Phi^{\dagger}\hat{z}$. Then $e = 0_{\mathsf{np}\times 1}$ if and only if $r = 0_{\mathsf{np}\times 1}$. Moreover, when $e = 0_{\mathsf{np}\times 1}$, the vector x is recovered by $\hat{x} := \Phi^{\dagger}\hat{z}$.

Proof: Note that any nonzero q-sparse error e does not lie in $\mathcal{R}(\Phi)$ by Proposition 1. 3). Hence, $e \neq 0_{\mathsf{np}\times 1}$ is equivalent to the condition that $\hat{z} = \Phi x + e \notin \mathcal{R}(\Phi)$. Since $\Phi\Phi^{\dagger}$ is a projection matrix and it projects $\hat{z}$ onto $\mathcal{R}(\Phi)$, we have $\hat{z} \notin \mathcal{R}(\Phi)$ if and only if $\hat{z} \neq \Phi\Phi^{\dagger}\hat{z}$. ■

Inspired by the error detection scheme for the noiseless case of Lemma 1, let us now consider a scheme for the case when the bounded noise $v \in \mathbb{R}^{\mathsf{np}}$ corrupts the measurements. For this, let

$$\rho_{\mathsf{p},\mathsf{q}}(\Phi) := \min\left\{\sigma_{\min}\left(\Phi_{\Lambda^{\mathsf{n}}}\right) : \Lambda \subset [\mathsf{p}],\ |\Lambda| = \mathsf{p}-\mathsf{q}\right\}$$

$$\eta_{\mathsf{p},\mathsf{q}}(\Phi) := \max\Big\{\left\|\Phi_{\Gamma_{\mathsf{i}}^{\mathsf{n}}}\left(\Phi_{\Lambda^{\mathsf{n}}}\right)^{\dagger}\right\|_2 : \mathsf{i} \in [\mathsf{p}]\backslash\Lambda,\ \Lambda \subset [\mathsf{p}],\ |\Lambda| = \mathsf{p}-\mathsf{q}\Big\}$$

$$\kappa_{\mathsf{p},\mathsf{q}}^{d}(\Phi) := (\sqrt{\mathsf{p}}+1)\sqrt{\mathsf{p}-\mathsf{q}}/\rho_{\mathsf{p},\mathsf{q}}(\Phi)$$

$$\kappa_{\mathsf{p},\mathsf{q}}^{e}(\Phi) := \left(\eta_{\mathsf{p},\mathsf{q}}(\Phi)\sqrt{\mathsf{p}-\mathsf{q}}+1\right)(\sqrt{\mathsf{p}}+1).$$

Then, the following theorem says that one can "practically" detect the q-sparse error in the noisy situation with the residual r given in (3).

Theorem 1: For the measurement $\hat{z} = \Phi x + v + e$ where $\Phi \in \mathbb{R}^{\mathsf{np}\times\mathsf{n}}$ is (n-stacked) q-error detectable, $x \in \mathbb{R}^{\mathsf{n}}$, $e \in \Sigma_{\mathsf{q}}^{\mathsf{n}}$, and $v \in \mathbb{R}^{\mathsf{np}}$ satisfying $\|v_{\mathsf{i}}^{\mathsf{n}}\|_2 \leq v_{\max}, {}^{\forall}\mathsf{i} \in [\mathsf{p}]$, let $\hat{x} = \Phi^{\dagger}\hat{z}$ and $r = \hat{z} - \Phi\hat{x}$. Then, the following hold:

1) $e \neq 0_{\mathsf{np}\times 1}$ if

$$\|r_{\mathsf{i}}^{\mathsf{n}}\|_2 = \|\hat{z}_{\mathsf{i}}^{\mathsf{n}} - \Phi_{\Gamma_{\mathsf{i}}^{\mathsf{n}}}^{\pi}\hat{x}\|_2 > \sqrt{\mathsf{p}}\, v_{\max} \quad \text{for some } \mathsf{i} \in [\mathsf{p}].$$

2) $\|e_{\mathsf{i}}^{\mathsf{n}}\|_2 \leq \kappa_{\mathsf{p},\mathsf{q}}^{e}(\Phi)v_{\max}, {}^{\forall}\mathsf{i} \in [\mathsf{p}]$, if

$$\|r_{\mathsf{i}}^{\mathsf{n}}\|_2 = \|\hat{z}_{\mathsf{i}}^{\mathsf{n}} - \Phi_{\Gamma_{\mathsf{i}}^{\mathsf{n}}}^{\pi}\hat{x}\|_2 \leq \sqrt{\mathsf{p}}\, v_{\max} \quad \text{for all } \mathsf{i} \in [\mathsf{p}].$$

In the case of 2), $\|\hat{x} - x\|_2 \leq \kappa_{\mathsf{p},\mathsf{q}}^{d}(\Phi)v_{\max}$.

Proof: The proof is omitted due to space limitations and can be found in [18]. ■

In fact, when the magnitude of e is small, one cannot differentiate between the noise v and the error e. Theorem 1. 2) reflects this fact and guarantees that the estimation error is small and $\hat{x}$ approximately estimates x.

B. Error Correctability and Reconstruction Scheme

In the noiseless case, the following notion of *error correctability* is introduced and characterized in this subsection.

Definition 2: A coding matrix $\Phi \in \mathbb{R}^{\mathsf{np}\times\mathsf{n}}$ is said to be *(n-stacked) q-error correctable* if, for all $x_1, x_2 \in \mathbb{R}^{\mathsf{n}}$ and $e_1, e_2 \in \Sigma_{\mathsf{q}}^{\mathsf{n}}$ such that $\Phi x_1 + e_1 = \Phi x_2 + e_2$, it holds that $x_1 = x_2$.

Now, one can easily obtain the following equivalence between the error correctability and the error detectability. The following proposition implies that one can detect twice the number of errors that can be corrected and reconstructed.

Proposition 2: The following are equivalent:

1) The matrix $\Phi \in \mathbb{R}^{\mathsf{np}\times\mathsf{n}}$ is (n-stacked) q-error correctable.
2) The matrix $\Phi \in \mathbb{R}^{\mathsf{np}\times\mathsf{n}}$ is (n-stacked) $2\mathsf{q}$-error detectable.

Proof: 1) ⇒ 2): Assume that $x, x' \in \mathbb{R}^{\mathsf{n}}$ and $e \in \Sigma_{2\mathsf{q}}^{\mathsf{n}}$ satisfying $\Phi x + e = \Phi x'$ are given. Let e_1 and e_2 be such that $e = e_1 - e_2$, where $e_1, e_2 \in \Sigma_{\mathsf{q}}^{\mathsf{n}}$. Thus, we have $\Phi x + e_1 = \Phi x' + e_2$. Since $\Phi \in \mathbb{R}^{\mathsf{np}\times\mathsf{n}}$ is q-error correctable, it follows that $x = x'$.

2) ⇒ 1): Assume that $x_1, x_2 \in \mathbb{R}^{\mathsf{n}}$ and $e_1, e_2 \in \Sigma_{\mathsf{q}}^{\mathsf{n}}$ satisfying $\Phi x_1 + e_1 = \Phi x_2 + e_2$ are given. Then, we have $\Phi x_1 + e = \Phi x_2$, where $e = e_1 - e_2 \in \Sigma_{2\mathsf{q}}^{\mathsf{n}}$. Since $\Phi \in \mathbb{R}^{\mathsf{np}\times\mathsf{n}}$ is $2\mathsf{q}$-error detectable, it follows that $x_1 = x_2$. ■

Based on the notion of q-error correctability, we discuss the problem of constructing a decoder that can actually correct (n-stacked) q errors and recover the original state x when $v = 0_{\mathsf{np}\times 1}$ in (2). That is, we find a map $\mathcal{D} : \mathbb{R}^{\mathsf{np}} \to \mathbb{R}^{\mathsf{n}}$ such that $\mathcal{D}(\hat{z}) = x$ where $\hat{z} = \Phi x + e \in \mathbb{R}^{\mathsf{np}}$ and $e \in \Sigma_{\mathsf{q}}^{\mathsf{n}}$. This is basically achieved through ℓ_0 minimization [19, Section 3]. Here we claim that searching over a finite set is enough to solve the minimization problem.

Theorem 2: For the measurement $\hat{z} = \Phi x + e \in \mathbb{R}^{\mathsf{np}}$ with (n-stacked) q-error correctable $\Phi \in \mathbb{R}^{\mathsf{np}\times\mathsf{n}}$, $x \in \mathbb{R}^{\mathsf{n}}$, and $e \in \Sigma_{\mathsf{q}}^{\mathsf{n}}$, it follows that

$$x = \arg\min_{\chi \in \mathcal{F}_{\mathsf{p},\mathsf{r}}(\hat{z})} \|\hat{z} - \Phi\chi\|_{0^{\mathsf{n}}} \tag{4}$$

$$= \arg\min_{\chi \in \mathcal{F}_{\mathsf{p},\mathsf{r}}(\hat{z})} \left|\left\{\mathsf{i} \in [\mathsf{p}] : \|\hat{z}_{\mathsf{i}}^{\mathsf{n}} - \Phi_{\Gamma_{\mathsf{i}}^{\mathsf{n}}}^{\pi}\chi\|_2 > 0\right\}\right| \tag{4'}$$

where

$$\mathcal{F}_{\mathsf{p},\mathsf{r}}(\hat{z}) := \left\{\left(\Phi_{\Lambda^{\mathsf{n}}}\right)^{\dagger}\hat{z}_{\Lambda^{\mathsf{n}}} \in \mathbb{R}^{\mathsf{n}} : \Lambda \subset [\mathsf{p}],\ |\Lambda| = \mathsf{p}-\mathsf{r}\right\}$$

and r is any integer satisfying $\mathsf{q} \leq \mathsf{r} \leq 2\mathsf{q}$.

Proof: We first show that the vector x belongs to $\mathcal{F}_{\mathsf{p},\mathsf{r}}(\hat{z})$. Pick any subset $\Lambda \subset (\mathsf{supp}^{\mathsf{n}}(e))^{c}$ satisfying $|\Lambda| = \mathsf{p}-\mathsf{r}$. Because $\Phi_{\Lambda^{\mathsf{n}}}$ has full column rank by Propositions 1. 2) and 2, it follows that $\chi = (\Phi_{\Lambda^{\mathsf{n}}})^{\dagger}\hat{z}_{\Lambda^{\mathsf{n}}} = (\Phi_{\Lambda^{\mathsf{n}}})^{\dagger}\Phi_{\Lambda^{\mathsf{n}}}x = x$. Hence, $x \in \mathcal{F}_{\mathsf{p},\mathsf{r}}(\hat{z})$. Now, it suffices to show that x is a minimizer of $\|\hat{z} - \Phi\chi\|_{0^{\mathsf{n}}}$. Suppose, for the sake of contradiction, that there exists $x' \neq x$ in $\mathcal{F}_{\mathsf{p},\mathsf{r}}(\hat{z})$ that minimizes $\|\hat{z} - \Phi\chi\|_{0^{\mathsf{n}}}$, then, with $e' := \hat{z} - \Phi x'$, we have that $\hat{z} = \Phi x' + e' = \Phi x + e$ and $\|e'\|_{0^{\mathsf{n}}} \leq \|e\|_{0^{\mathsf{n}}} \leq \mathsf{q}$ because e' is a minimal solution. This contradicts the assumption that Φ is q-error correctable. ■

This theorem claims that it is enough to search over the finite set $\mathcal{F}_{\mathsf{p},\mathsf{r}}(\hat{z})$, not the whole space $\mathbb{R}^{\mathsf{n}}$, to solve (4). Keeping in mind the fact that $|\mathcal{F}_{\mathsf{p},\mathsf{r}}(\hat{z})| \leq \binom{\mathsf{p}}{\mathsf{p}-\mathsf{r}} = \binom{\mathsf{p}}{\mathsf{r}}$, one can choose any integer r between q and $2\mathsf{q}$ to minimize $\binom{\mathsf{p}}{\mathsf{r}}$.

Remark 2: The ℓ_0 minimization problem over $\mathbb{R}^{\mathsf{n}}$ is shown to be NP-hard [20]. Whereas previous research efforts have been devoted to a relaxation of the problem by imposing some additional conditions (e.g., [9], [11]), Theorem 2 actually relieves the computational complexity by reducing the search space to a "finite" set. It is a kind of combinatorial approach that tests only $\binom{\mathsf{p}}{\mathsf{r}} \leq \mathsf{p}^{\mathsf{r}}$ (or $\binom{\mathsf{p}}{\mathsf{p}-\mathsf{r}} \leq \mathsf{p}^{\mathsf{p}-\mathsf{r}}$) candidates with the freedom of selecting r between q and $2\mathsf{q}$, whereas naive brute-force search algorithm without any information on error correctability has no choice but to test all $\binom{\mathsf{p}}{1} + \binom{\mathsf{p}}{2} + \ldots + \binom{\mathsf{p}}{\mathsf{p}} \approx 2^{\mathsf{p}}$ combinations. In our case, the computational efforts decrease drastically by selecting $\mathsf{r} = \mathsf{q}$

when $\mathsf{q} \ll \mathsf{p}$ (or selecting $\mathsf{r} = 2\mathsf{q}$ when $\mathsf{q} \approx \mathsf{p}/2$) for example. Compared with other combinatorial algorithms in [1], [4], [12], Theorem 2 is more relaxed by introducing r that can vary between q and $2\mathsf{q}$. ♦

Finally, the following lemma presents a simple criterion to verify whether a given vector $\hat{x} \in \mathbb{R}^{\mathsf{n}}$ coincides with the original input x.

Lemma 2: For the measurement $\hat{z} = \Phi x + e \in \mathbb{R}^{\mathsf{np}}$ with (n-stacked) q-error correctable $\Phi \in \mathbb{R}^{\mathsf{np}\times\mathsf{n}}$, $x \in \mathbb{R}^{\mathsf{n}}$, and $e \in \Sigma_{\mathsf{q}}^{\mathsf{n}}$,

$$\|\hat{z} - \Phi\hat{x}\|_{0^{\mathsf{n}}} \le \mathsf{q} \quad \text{if and only if} \quad x = \hat{x}.$$

Proof: (if): This is trivial because $\|\hat{z} - \Phi\hat{x}\|_{0^{\mathsf{n}}} = \|e\|_{0^{\mathsf{n}}} \le \mathsf{q}$. (only if): Define $\hat{e} := \hat{z} - \Phi\hat{x}$, then $\hat{z} = \Phi\hat{x} + \hat{e} = \Phi x + e$, where $e, \hat{e} \in \Sigma_{\mathsf{q}}^{\mathsf{n}}$. Since Φ is q-error correctable, it follows from Definition 2 that $x = \hat{x}$. ■

Now, bounded noise $v \in \mathbb{R}^{\mathsf{np}}$ satisfying $\|v_{\mathsf{i}}^{\mathsf{n}}\|_2 \le v_{\max}$ for all $\mathsf{i} \in [\mathsf{p}]$ is taken into account and a state recovery scheme estimating x is presented. More precisely, we show that any solution (χ^*, ε^*) to the following relaxed ℓ_0 minimization problem yields an approximation of x as $\hat{x} = \chi^*$:

$$\begin{aligned} \min_{\chi \in \mathcal{F}_{\mathsf{p},\mathsf{r}}(\hat{z}),\, \varepsilon \in \mathbb{R}^{\mathsf{np}}} \quad & \|\varepsilon\|_{0^{\mathsf{n}}} \\ \text{subject to} \quad & \|\hat{z}_{\mathsf{i}}^{\mathsf{n}} - \Phi_{\Gamma_{\mathsf{i}}^{\mathsf{n}}}^{\pi}\chi - \varepsilon_{\mathsf{i}}^{\mathsf{n}}\|_2 \le v'_{\max},\ {}^{\forall}\mathsf{i} \in [\mathsf{p}] \end{aligned} \tag{5}$$

where r is any integer satisfying $\mathsf{q} \le \mathsf{r} \le 2\mathsf{q}$ and

$$\begin{aligned} v'_{\max} &:= \vartheta_{\mathsf{p},\mathsf{q},\mathsf{r}}(\Phi) v_{\max} \\ &:= \max\left\{\eta'_{\mathsf{p},\mathsf{q},\mathsf{r}}(\Phi)\sqrt{\mathsf{p}-\mathsf{r}} + 1, \sqrt{\mathsf{p}-\mathsf{r}}\right\} v_{\max} \\ \eta'_{\mathsf{p},\mathsf{q},\mathsf{r}}(\Phi) &:= \max_{\substack{\Lambda \subset [\mathsf{p}] \\ |\Lambda| = \mathsf{p}-\mathsf{q}}} \min_{\substack{\bar{\Lambda} \subset \Lambda \\ |\bar{\Lambda}| = \mathsf{p}-\mathsf{r}}} \max_{\mathsf{i} \in \Lambda \setminus \bar{\Lambda}} \left\| \Phi_{\Gamma_{\mathsf{i}}^{\mathsf{n}}} \left(\Phi_{\bar{\Lambda}^{\mathsf{n}}}\right)^{\dagger} \right\|_2. \end{aligned}$$

The above optimization problem is not easily implementable because the variable ε is searched over $\mathbb{R}^{\mathsf{np}}$ under constraints. Hence, we present another optimization problem, which may be considered as a relaxation of (4′):

$$\hat{x} = \arg\min_{\chi \in \mathcal{F}_{\mathsf{p},\mathsf{r}}(\hat{z})} \left|\left\{\mathsf{i} \in [\mathsf{p}] : \|\hat{z}_{\mathsf{i}}^{\mathsf{n}} - \Phi_{\Gamma_{\mathsf{i}}^{\mathsf{n}}}^{\pi}\chi\|_2 > v'_{\max}\right\}\right|. \tag{5′}$$

Whereas the problem (5) or (5′) need not have a unique solution, the following theorem shows equivalence between (5) and (5′), and presents an upper bound of $\|\hat{x} - x\|_2$ for any solution $\hat{x}$ of (5) or (5′).

Theorem 3: For the measurement $\hat{z} = \Phi x + e + v \in \mathbb{R}^{\mathsf{np}}$ with (n-stacked) q-error correctable $\Phi \in \mathbb{R}^{\mathsf{np}\times\mathsf{n}}$, $x \in \mathbb{R}^{\mathsf{n}}$, $e \in \Sigma_{\mathsf{q}}^{\mathsf{n}}$, and $v \in \mathbb{R}^{\mathsf{np}}$ such that $\|v_{\mathsf{i}}^{\mathsf{n}}\|_2 \le v_{\max}$, ${}^{\forall}\mathsf{i} \in [\mathsf{p}]$, the following hold:

1) Two optimization problems (5) and (5′) are equivalent (that is, a solution $\hat{x}$ to (5) is also a solution to (5′) and vice versa).
2) For any solution $\hat{x}$, $\|\hat{x} - x\|_2 \le \kappa^{c}_{\mathsf{p},\mathsf{q},\mathsf{r}}(\Phi)\, v_{\max}$, where

$$\kappa^{c}_{\mathsf{p},\mathsf{q},\mathsf{r}}(\Phi) := (\vartheta_{\mathsf{p},\mathsf{q},\mathsf{r}}(\Phi) + 1)\sqrt{\mathsf{p} - 2\mathsf{q}}/\rho_{\mathsf{p},2\mathsf{q}}(\Phi).$$

Proof: The proof is omitted due to space limitations and can be found in [18]. ■

As in Lemma 2, a simple criterion to check whether a given vector $\hat{x} \in \mathbb{R}^{\mathsf{n}}$ is close to the original x with noisy measurements, is also derived in the following theorem.

Theorem 4: For the measurement $\hat{z} = \Phi x + e + v \in \mathbb{R}^{\mathsf{np}}$ with (n-stacked) q-error correctable $\Phi \in \mathbb{R}^{\mathsf{np}\times\mathsf{n}}$, $x \in \mathbb{R}^{\mathsf{n}}$, $e \in \Sigma_{\mathsf{q}}^{\mathsf{n}}$, and $v \in \mathbb{R}^{\mathsf{np}}$ such that $\|v_{\mathsf{i}}^{\mathsf{n}}\|_2 \le v_{\max}$, ${}^{\forall}\mathsf{i} \in [\mathsf{p}]$, the following hold:

1) $\|\hat{x} - x\|_2 \le \kappa^{c}_{\mathsf{p},\mathsf{q},\mathsf{r}}(\Phi)\, v_{\max}$ if $\hat{x}$ satisfies

$$\left|\left\{\mathsf{i} \in [\mathsf{p}] : \|\hat{z}_{\mathsf{i}}^{\mathsf{n}} - \Phi_{\Gamma_{\mathsf{i}}^{\mathsf{n}}}^{\pi}\hat{x}\|_2 > v'_{\max}\right\}\right| \le \mathsf{q}.$$

2) $\|\hat{x} - x\|_2 > \kappa^{c'}_{\mathsf{p},\mathsf{q},\mathsf{r}}(\Phi)\, v_{\max}$ if $\hat{x}$ satisfies

$$\left|\left\{\mathsf{i} \in [\mathsf{p}] : \|\hat{z}_{\mathsf{i}}^{\mathsf{n}} - \Phi_{\Gamma_{\mathsf{i}}^{\mathsf{n}}}^{\pi}\hat{x}\|_2 > v'_{\max}\right\}\right| > \mathsf{q}$$

where $\kappa^{c'}_{\mathsf{p},\mathsf{q},\mathsf{r}}(\Phi) := (\vartheta_{\mathsf{p},\mathsf{q},\mathsf{r}}(\Phi) - 1)/\max_{\mathsf{i}\in[\mathsf{p}]} \|\Phi_{\Gamma_{\mathsf{i}}^{\mathsf{n}}}\|_2$.

Proof: The proof is omitted due to space limitations and can be found in [18]. ■

III. Characterization of Redundant Observability

In this section, we introduce the *redundant observability* and relate that concept to the *dynamic security index*, *attack detectability*, and *observability under sensor attacks*. It will soon be revealed that an observability matrix behaves in the same way as a coding matrix as examined in the previous section, and hence its properties determine resilience of control systems under sensor attacks.

A. Redundant Observability

From a control theoretical viewpoint, the notion of *redundant observability* for the system (1) is defined as follows.

Definition 3: The pair (A, C) or the dynamical system (1) is said to be *q-redundant observable* if the pair (A, C_{Λ}^{π}) is observable for any $\Lambda \subset [\mathsf{p}]$ satisfying $|\Lambda| \ge \mathsf{p} - \mathsf{q}$.

To characterize the redundant observability in the following proposition, we first obtain the observability matrix $G \in \mathbb{R}^{\mathsf{np}\times\mathsf{n}}$ as follows:

$$G := \left[G_1^{\top}\, G_2^{\top}\, \cdots\, G_{\mathsf{p}}^{\top}\right]^{\top} \tag{6}$$

where $G_{\mathsf{i}} := [c_{\mathsf{i}}^{\top}\, (c_{\mathsf{i}}A)^{\top}\, \cdots\, (c_{\mathsf{i}}A^{\mathsf{n}-1})^{\top}]^{\top}$ is an observability matrix of the pair (A, c_{i}).

Proposition 3: The following are equivalent:

1) The pair (A, C) is q-redundant observable.
2) The matrix G is (n-stacked) q-error detectable.

Proof: From the fact that $G_{\Lambda^{\mathsf{n}}}^{\pi}$ is the observability matrix of the pair (A, C_{Λ}^{π}), the pair (A, C) is q-redundant observable if and only if $G_{\Lambda^{\mathsf{n}}}^{\pi}$ has full column rank for any $\Lambda \subset [\mathsf{p}]$ satisfying $|\Lambda| \ge \mathsf{p} - \mathsf{q}$. Thus, the result directly follows from Proposition 1. ■

B. Attack Detectability and Dynamic Security Index

Assume tentatively that there is no control input, disturbance, nor noise in the system (1) so that we can focus on the attack signal only. Then, the output measurements for a finite time period are collected and the stacked output sequence is computed as

$$\begin{aligned} \bar{y}^{[0:\mathsf{n}-1]} &:= \begin{bmatrix} \bar{y}_1^{[0:\mathsf{n}-1]} \\ \bar{y}_2^{[0:\mathsf{n}-1]} \\ \vdots \\ \bar{y}_{\mathsf{p}}^{[0:\mathsf{n}-1]} \end{bmatrix} = \begin{bmatrix} G_1 \\ G_2 \\ \vdots \\ G_{\mathsf{p}} \end{bmatrix} x(0) + \begin{bmatrix} a_1^{[0:\mathsf{n}-1]} \\ a_2^{[0:\mathsf{n}-1]} \\ \vdots \\ a_{\mathsf{p}}^{[0:\mathsf{n}-1]} \end{bmatrix} \\ &= Gx(0) + a^{[0:\mathsf{n}-1]} \end{aligned} \tag{7}$$

where $\bar{y}_{\mathsf{i}}^{[0:\mathsf{n}-1]} := [\bar{y}_{\mathsf{i}}(0)\, \bar{y}_{\mathsf{i}}(1)\, \cdots\, \bar{y}_{\mathsf{i}}(\mathsf{n}-1)]^{\top}$ and $a_{\mathsf{i}}^{[0:\mathsf{n}-1]} := [a_{\mathsf{i}}(0)\, a_{\mathsf{i}}(1)\, \cdots\, a_{\mathsf{i}}(\mathsf{n}-1)]^{\top}$. Noting that the situation is exactly the same as the noiseless case in Section II-A and $a^{[0:\mathsf{n}-1]}$ is (n-stacked) q-sparse by Assumption 2, we can introduce the notion of *q-attack detectability* of the system (1) as follows.

Definition 4: The pair (A, C) or the dynamical system (1) without disturbances/noises is said to be *q-attack detectable* if, for all $x(0), x'(0) \in \mathbb{R}^{\mathsf{n}}$ and $a^{[0:\mathsf{n}-1]} \in \Sigma_{\mathsf{q}}^{\mathsf{n}}$ such that $Gx(0) + a^{[0:\mathsf{n}-1]} = Gx'(0)$, it holds that $x(0) = x'(0)$.

Furthermore, a direct comparison between Definitions 1 and 4 simply leads to the following proposition.

Proposition 4: The following are equivalent:

1) The pair (A, C) is q-attack detectable.
2) The matrix G is (n-stacked) q-error detectable.

As a tool for the vulnerability analysis of a system, the security index quantifies fundamental limitations on the attack detectability. That is, the *dynamic security index* of the system (1), $\alpha_d(A,C)$, is defined by the minimum number of sensor attacks for adversaries to remain undetectable and is computed by examining the system's strong observability in [8] as

$$\alpha_d(A,C) = \min_{v\in\mathcal{V}(A)} \|Cv\|_0. \tag{8}$$

It is shown in the following proposition that the dynamic security index can also be characterized by the error detectability of the observability matrix G through its cospark.

Proposition 5: For $\alpha_d(A,C)$ given in (8), it holds that

$$\alpha_d(A,C) = \min_{x\in\mathbb{R}^{\mathsf{n}},\, x\neq 0_{\mathsf{n}\times 1}} \|Gx\|_{0^{\mathsf{n}}} = \mathsf{cospark}^{\mathsf{n}}(G). \tag{9}$$

Proof: When $Av=\lambda v$, one can trivially check that

$$\min_{v\in\mathcal{V}(A)} \|Cv\|_0 = \min_{v\in\mathcal{V}(A)} \|Gv\|_{0^{\mathsf{n}}}$$

since $G_{\mathsf{i}}v = [c_{\mathsf{i}}v\ \lambda c_{\mathsf{i}}v\ \cdots\ \lambda^{\mathsf{n}-1}c_{\mathsf{i}}v]^\top$. Noting that

$$\min_{x\in\mathbb{C}^{\mathsf{n}},\, x\neq 0_{\mathsf{n}\times 1}} \|Gx\|_{0^{\mathsf{n}}} = \min_{x\in\mathbb{R}^{\mathsf{n}},\, x\neq 0_{\mathsf{n}\times 1}} \|Gx\|_{0^{\mathsf{n}}}$$

because $G\in\mathbb{R}^{\mathsf{np}\times\mathsf{n}}$ is a real matrix, it suffices to show that

$$\min_{v\in\mathcal{V}(A)} \|Gv\|_{0^{\mathsf{n}}} = \min_{x\in\mathbb{C}^{\mathsf{n}},\, x\neq 0_{\mathsf{n}\times 1}} \|Gx\|_{0^{\mathsf{n}}}.$$

Now, we claim that there exists $v^*\in\mathcal{V}(A)$ such that

$$\|Gv^*\|_{0^{\mathsf{n}}} = \min_{x\in\mathbb{C}^{\mathsf{n}},\, x\neq 0_{\mathsf{n}\times 1}} \|Gx\|_{0^{\mathsf{n}}}.$$

Let us denote the optimal value of the problem (9) by

$$\alpha^* := \min_{x\in\mathbb{R}^{\mathsf{n}},\, x\neq 0_{\mathsf{n}\times 1}} \|Gx\|_{0^{\mathsf{n}}}.$$

By the equivalence between Proposition 1. 2) and 3) with the observability matrix G, there exists an index set $\Lambda\subset[\mathsf{p}]$ satisfying $|\Lambda| = \mathsf{p}-\alpha^*$ such that the observability matrix $G^{\pi}_{\Lambda^{\mathsf{n}}}$ does not have full column rank but the observability matrix $G^{\pi}_{(\Lambda\cup\{\mathsf{i}\})^{\mathsf{n}}}$ has full column rank for every $\mathsf{i}\in\Lambda^c$. That is, the pair (A, C^{π}_{Λ}) is not observable but the pair $(A, C^{\pi}_{\Lambda\cup\{\mathsf{i}\}})$ is observable for every $\mathsf{i}\in\Lambda^c$. Applying the Popov–Belevitch–Hautus (PBH) observability test, we conclude that there exist $\lambda^*\in\mathbb{C}$ and $v^*\in\mathcal{V}(A)$ such that

$$\begin{bmatrix} \lambda^* I_{\mathsf{n}\times\mathsf{n}} - A \\ C^{\pi}_{\Lambda} \end{bmatrix} v^* = \begin{bmatrix} 0_{\mathsf{n}\times 1} \\ 0_{(\mathsf{p}-\alpha^*)\times 1} \end{bmatrix} \text{ and } c_{\mathsf{i}}v^*\neq 0,\ {}^{\forall}\mathsf{i}\in\Lambda^c.$$

The claim easily follows by verifying that $\|Gv^*\|_{0^{\mathsf{n}}} = \alpha^*$. ■

C. Observability Under Sparse Sensor Attacks

In this subsection, the notion of *observability under* q*-sparse sensor attacks* is introduced and an equivalent condition is directly derived from the definition as follows.

Definition 5: The pair (A,C) or the dynamical system (1) without disturbances/noises is said to be *observable under* q*-sparse sensor attacks* if the initial state $x(0)$ can be determined from the output $\bar{y}$ over a finite number of sampling steps with any sensor attack a satisfying Assumption 2.

Proposition 6: The following are equivalent:

1) The pair (A,C) is observable under q-sparse sensor attacks.
2) The matrix G is (n-stacked) q-error correctable.

IV. Design of Attack-Resilient Estimator

An attack-resilient state estimator $\mathcal{E}$, which combines the partial observers $\mathcal{O}_{\mathsf{i}}$ and the decoder $\mathcal{D}$, is designed in this section. First, the partial observers $\mathcal{O}_{\mathsf{i}}$ are designed by applying the Kalman observability decomposition to each sensor output. Second, the previously developed error correction technique tailored into this specific problem constitutes the decoder $\mathcal{D}$ and it recovers the original state variable x.

A. Design of Partial Observers

With only one measurement $\bar{y}_{\mathsf{i}}(k)$ of the plant (1), a single-output system is obtained as follows:

$$\mathcal{P}_{\mathsf{i}}: \begin{cases} x(k+1) = Ax(k) + Bu(k) + d(k) \\ \bar{y}_{\mathsf{i}}(k) = c_{\mathsf{i}}x(k) + n_{\mathsf{i}}(k) + a_{\mathsf{i}}(k). \end{cases} \tag{10}$$

The observability matrix G_{i} of (10) is used to divide the n-dimensional state space into two subspaces. To derive a transformation matrix, first, let ν_{i} be the observability index of (A,c_{i}), i.e., $\nu_{\mathsf{i}} := \mathsf{rank}(G_{\mathsf{i}})$. Then the set of the first ν_{i} rows of G_{i} is linearly independent. The null space of G_{i}, $\mathcal{N}(G_{\mathsf{i}})$, which is A-invariant, is the unobservable subspace. Furthermore, the quotient space $\mathbb{R}^{\mathsf{n}}/\mathcal{N}(G_{\mathsf{i}})$ is sometimes called, with abuse of terminology, the observable subspace. The matrices $Z_{\mathsf{i}}\in\mathbb{R}^{\mathsf{n}\times\nu_{\mathsf{i}}}$ and $W_{\mathsf{i}}\in\mathbb{R}^{\mathsf{n}\times(\mathsf{n}-\nu_{\mathsf{i}})}$ are selected such that their columns are orthonormal bases of $\mathcal{N}(G_{\mathsf{i}})^{\perp}(=\mathcal{R}(G_{\mathsf{i}}^{\top}))$ and $\mathcal{N}(G_{\mathsf{i}})$, respectively. Finally, by the Kalman observability decomposition, the state x is decomposed into the observable substate $z_{\mathsf{i}}\in\mathbb{R}^{\nu_{\mathsf{i}}}$ and the unobservable substate $w_{\mathsf{i}}\in\mathbb{R}^{\mathsf{n}-\nu_{\mathsf{i}}}$ with a similarity transformation

$$\begin{bmatrix} z_{\mathsf{i}}^\top & w_{\mathsf{i}}^\top \end{bmatrix}^\top = \begin{bmatrix} Z_{\mathsf{i}} & W_{\mathsf{i}} \end{bmatrix}^\top x. \tag{11}$$

Since it follows that $Z_{\mathsf{i}}^\top A W_{\mathsf{i}} = O_{\nu_{\mathsf{i}}\times(\mathsf{n}-\nu_{\mathsf{i}})}$ and $c_{\mathsf{i}}W_{\mathsf{i}} = 0_{1\times(\mathsf{n}-\nu_{\mathsf{i}})}$ from the construction of Z_{i} and W_{i}, the change of variable (11) leads the original single-output system (10) to the decomposed form of

$$\mathcal{P}'_{\mathsf{i}}: \begin{cases} \begin{bmatrix} z_{\mathsf{i}}(k+1) \\ w_{\mathsf{i}}(k+1) \end{bmatrix} = \begin{bmatrix} Z_{\mathsf{i}}^\top A Z_{\mathsf{i}} & O_{\nu_{\mathsf{i}}\times(\mathsf{n}-\nu_{\mathsf{i}})} \\ W_{\mathsf{i}}^\top A Z_{\mathsf{i}} & W_{\mathsf{i}}^\top A W_{\mathsf{i}} \end{bmatrix} \begin{bmatrix} z_{\mathsf{i}}(k) \\ w_{\mathsf{i}}(k) \end{bmatrix} \\ \qquad\qquad + \begin{bmatrix} Z_{\mathsf{i}}^\top B \\ W_{\mathsf{i}}^\top B \end{bmatrix} u(k) + \begin{bmatrix} Z_{\mathsf{i}}^\top \\ W_{\mathsf{i}}^\top \end{bmatrix} d(k) \\ \bar{y}_{\mathsf{i}}(k) = \begin{bmatrix} c_{\mathsf{i}}Z_{\mathsf{i}} & 0_{1\times(\mathsf{n}-\nu_{\mathsf{i}})} \end{bmatrix} \begin{bmatrix} z_{\mathsf{i}}(k) \\ w_{\mathsf{i}}(k) \end{bmatrix} + n_{\mathsf{i}}(k) + a_{\mathsf{i}}(k). \end{cases} \tag{12}$$

By dropping the unobservable substate w_{i} from (12), the observable quotient subsystem of (12) is obtained as

$$\mathcal{P}^o_{\mathsf{i}}: \begin{cases} z_{\mathsf{i}}(k+1) = S_{\mathsf{i}}z_{\mathsf{i}}(k) + Z_{\mathsf{i}}^\top B u(k) + Z_{\mathsf{i}}^\top d(k) \\ \bar{y}_{\mathsf{i}}(k) = t_{\mathsf{i}}z_{\mathsf{i}}(k) + n_{\mathsf{i}}(k) + a_{\mathsf{i}}(k) \end{cases} \tag{13}$$

where $S_{\mathsf{i}} := Z_{\mathsf{i}}^\top A Z_{\mathsf{i}}$, $t_{\mathsf{i}} := c_{\mathsf{i}}Z_{\mathsf{i}}$ and the pair $(S_{\mathsf{i}}, t_{\mathsf{i}})$ is observable.

Then, the partial observer $\mathcal{O}_{\mathsf{i}}$ is designed by a Luenberger observer for (13) given in the following form:

$$\mathcal{O}_{\mathsf{i}}:\ \hat{z}_{\mathsf{i}}(k+1) = F_{\mathsf{i}}\hat{z}_{\mathsf{i}}(k) + Z_{\mathsf{i}}^\top B u(k) + L_{\mathsf{i}}\bar{y}_{\mathsf{i}}(k) \tag{14}$$

where the injection gain L_{i} is chosen so that $F_{\mathsf{i}} := S_{\mathsf{i}} - L_{\mathsf{i}}t_{\mathsf{i}}$ is Schur stable. The dynamics of state estimation error $\tilde{z}_{\mathsf{i}} := \hat{z}_{\mathsf{i}} - z_{\mathsf{i}}$ is governed by

$$\mathcal{F}_{\mathsf{i}}: \tilde{z}_{\mathsf{i}}(k+1) = F_{\mathsf{i}}\tilde{z}_{\mathsf{i}}(k) + L_{\mathsf{i}}n_{\mathsf{i}}(k) - Z_{\mathsf{i}}^\top d(k) + L_{\mathsf{i}}a_{\mathsf{i}}(k)$$

whose solution becomes

$$\tilde{z}_{\mathsf{i}}(k) = v_{\mathsf{i}}(k) + e_{\mathsf{i}}(k) \tag{15}$$

where $v_{\mathsf{i}}(k) := F_{\mathsf{i}}^k \tilde{z}_{\mathsf{i}}(0) + \sum_{j=0}^{k-1} F_{\mathsf{i}}^{k-1-j}(L_{\mathsf{i}} n_{\mathsf{i}}(j) - Z_{\mathsf{i}}^\top d(j))$ and $e_{\mathsf{i}}(k) := \sum_{j=0}^{k-1} F_{\mathsf{i}}^{k-1-j} L_{\mathsf{i}} a_{\mathsf{i}}(j)$. Here, the attack-induced estimation error vector $e_{\mathsf{i}}(k)$ may have arbitrary values. For all $k \geq 0$ and $\mathsf{i} \in [\mathsf{p}]$, there exist $\mu_F \geq 1$ and $0 < \beta < 1$ such that $\|F_{\mathsf{i}}^k\|_2 \leq \mu_F \beta^k$ since F_{i} is Schur stable. In addition, for some μ_L and μ_Z, it holds that $\|F_{\mathsf{i}}^k L_{\mathsf{i}}\|_2 \leq \mu_L \beta^k$ and $\|F_{\mathsf{i}}^k Z_{\mathsf{i}}^\top\|_2 \leq \mu_Z \beta^k$. Then, one can easily show that

$$\|v_{\mathsf{i}}(k)\|_2 \leq \mu_F \|\tilde{z}_{\mathsf{i}}(0)\|_2 \beta^k + w_{\max} \leq v_{\max}(k) \tag{16}$$

where $w_{\max} := (\mu_L n_{\max} + \mu_Z d_{\max})/(1-\beta)$ and $v_{\max}(k) := \max_{\mathsf{i}\in[\mathsf{p}]}\{\mu_F \|\tilde{z}_{\mathsf{i}}(0)\|_2 \beta^k + w_{\max}\}$. As k increases, $v_{\max}(k)$ converges to $w_{\max}$.

B. Design of the Decoder

The decoder collects all the data $\hat{z}_{\mathsf{i}}$ from the partial observers $\mathcal{O}_{\mathsf{i}}$ and formulates the problem in the form of (2). To this end, $Z_{\mathsf{i}}^\top x = z_{\mathsf{i}}$ in (11), is used. Appending $\mathsf{n} - \nu_{\mathsf{i}}$ zero row vectors, $0_{1\times\mathsf{n}}$, to each $Z_{\mathsf{i}}^\top$ and stacking them all, we have

$$\begin{bmatrix} Z_1^\top \\ \vdots \\ Z_{\mathsf{p}}^\top \end{bmatrix} x(k) = \begin{bmatrix} z_1(k) \\ \vdots \\ z_{\mathsf{p}}(k) \end{bmatrix} = \begin{bmatrix} \hat{z}_1(k) \\ \vdots \\ \hat{z}_{\mathsf{p}}(k) \end{bmatrix} - \begin{bmatrix} \tilde{z}_1(k) \\ \vdots \\ \tilde{z}_{\mathsf{p}}(k) \end{bmatrix} \tag{17}$$

where

$$\begin{aligned} Z_{\mathsf{i}}^{\mathsf{n}\top} &:= \begin{bmatrix} Z_{\mathsf{i}}^\top \\ O_{(\mathsf{n}-\nu_{\mathsf{i}})\times\mathsf{n}} \end{bmatrix}, & z_{\mathsf{i}}^{\mathsf{n}}(k) &:= \begin{bmatrix} z_{\mathsf{i}}(k) \\ 0_{(\mathsf{n}-\nu_{\mathsf{i}})\times 1} \end{bmatrix}, \\ \hat{z}_{\mathsf{i}}^{\mathsf{n}}(k) &:= \begin{bmatrix} \hat{z}_{\mathsf{i}}(k) \\ 0_{(\mathsf{n}-\nu_{\mathsf{i}})\times 1} \end{bmatrix}, & \tilde{z}_{\mathsf{i}}^{\mathsf{n}}(k) &:= \begin{bmatrix} \tilde{z}_{\mathsf{i}}(k) \\ 0_{(\mathsf{n}-\nu_{\mathsf{i}})\times 1} \end{bmatrix}. \end{aligned} \tag{18}$$

This augmentation of zeros is to match the size of each matrix so that it agrees with the n-stacked vector considered in Section II. Now, (17) with (15) is written in a compact form as

$$\hat{z}(k) = \Phi x(k) + \tilde{z}(k) = \Phi x(k) + v(k) + e(k) \in \mathbb{R}^{\mathsf{np}} \tag{19}$$

where $\Phi := [Z_1^{\mathsf{n}}\, Z_2^{\mathsf{n}} \cdots Z_{\mathsf{p}}^{\mathsf{n}}]^\top$. It is supposed that additional zero elements are also appended to $v_{\mathsf{i}}(k)$ and $e_{\mathsf{i}}(k)$ as in (18).

Since (19) exactly matches with (2), one can directly apply the error correction technique developed in Section II into (19). Theorems 3 and 4 are mainly employed so as to recover $x(k)$. Before applying them, one should check that three conditions on those theorems are satisfied for the given system (1): boundedness of $v(k)$, q-sparsity of $e(k)$, and q-error correctability of Φ. The first two conditions are easily satisfied by Assumptions 1 and 2. That is, the noise vector $v_{\mathsf{i}}(k)$ is bounded by $v_{\max}(k)$ for all $\mathsf{i} \in [\mathsf{p}]$ by (16), which is induced from Assumption 1. Since the error vector $e_{\mathsf{i}}(k)$ depends only on the attack element $a_{\mathsf{i}}(j)$ for $0 \leq j \leq k-1$ and $a(j)$ is q-sparse according to Assumption 2, the vector $e(k)$ is (n-stacked) q-sparse. An additional assumption should be declared for the last condition, the q-error correctability of Φ, to be fulfilled.

Assumption 3: The pair (A, C) is $2\mathsf{q}$-redundant observable. ◊

Under Assumption 3, Propositions 2 and 3 ensure the q-error correctability of G where G is given in (6). Thus, Φ is also q-error correctable because $\mathcal{R}(\Phi_{\Lambda^{\mathsf{n}}}^\top) = \mathcal{R}(G_{\Lambda^{\mathsf{n}}}^\top)$ for any $\Lambda \subset [\mathsf{p}]$ by the construction of Φ and its elements $Z_{\mathsf{i}}^\top$. Therefore, all three conditions on Theorems 3 and 4 hold.

The decoder's configuration is sketched in Fig. 2 and its operation is described in Algorithm 1. During the operation of the decoder, the monitoring scheme that the selector $\mathcal{S}$ and the switch s perform is

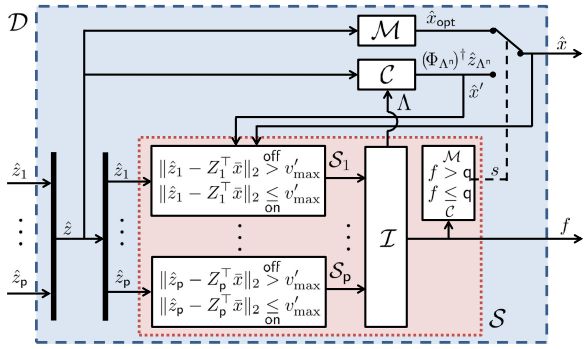

Fig. 2. Configuration of the decoder.

Algorithm 1: Operation of the decoder.

Input: $\hat{z}_1, \hat{z}_2, \ldots, \hat{z}_{\mathsf{p}}$
Output: $\hat{x}$, f
Initialization: $\Lambda = [\mathsf{p}]$
1: **while** System (1) is running **do**
2: $\hat{x}' = (\Phi_{\Lambda^{\mathsf{n}}})^\dagger \hat{z}_{\Lambda^{\mathsf{n}}}$
3: $f = \mathsf{p} - |\{\mathsf{i} \in [\mathsf{p}] : \|\hat{z}_{\mathsf{i}} - Z_{\mathsf{i}}^\top \hat{x}'\|_2 \leq v'_{\max}\}|$
4: **if** $f \leq \mathsf{q}$ **then**
5: Switch s selects the line from $\mathcal{C}$
6: Calculator $\mathcal{C}$ sets $\hat{x} = \hat{x}'$
7: **else**
8: Switch s selects the line from $\mathcal{M}$
9: Minimizer $\mathcal{M}$ solves (5′) and produces $\hat{x} = \hat{x}_{\mathsf{opt}}$
10: Update $\Lambda = \{\mathsf{i} \in [\mathsf{p}] : \|\hat{z}_{\mathsf{i}} - Z_{\mathsf{i}}^\top \hat{x}\|_2 \leq v'_{\max}\}$
11: **end if**
12: **end while**

running on the basis of Theorem 4, whereas the calculator $\mathcal{C}$ and the minimizer $\mathcal{M}$ have their roots in Theorems 1 and 3, respectively. Note that $\bar{x}$ of each selector $\mathcal{S}_{\mathsf{i}}$ in Fig. 2 represents $\hat{x}'$ at the stage of monitoring (line 3 of Algorithm 1) and $\hat{x}$ during the updating step (line 10 of Algorithm 1). If $f \leq \mathsf{q}$, the successful state estimation is ensured by Theorem 4. 1). More specifically, we have $\|\hat{x} - x\|_2 \leq \kappa^c_{\mathsf{p},\mathsf{q},\mathsf{r}}(\Phi) v_{\max}$. In this case, the index set Λ can be supposed to be attack-free, and hence the calculator $\mathcal{C}$ can recover the original state x approximately by $\hat{x} = (\Phi_{\Lambda^{\mathsf{n}}})^\dagger \hat{z}_{\Lambda^{\mathsf{n}}}$, which is attributed to Theorem 1. On the other hand, if $f > \mathsf{q}$, the state estimate $\hat{x}$ is not close enough to the original state x by Theorem 4. 2). Hence, the algorithm goes to the minimizer step (i.e., the switch s chooses the side of the minimizer $\mathcal{M}$) to figure out new healthy sensors and the state estimates $\hat{x}$ by $\hat{x}_{\mathsf{opt}}$. Furthermore, Theorem 3 guarantees that $\|\hat{x} - x\|_2 \leq \kappa^c_{\mathsf{p},\mathsf{q},\mathsf{r}}(\Phi) v_{\max}$. These results are summarized in the following theorem.

Theorem 5: Under Assumptions 1–3, the estimator $\mathcal{E}$ equipped with the observers $\mathcal{O}_{\mathsf{i}}$ given by (14) and the decoder $\mathcal{D}$ employing Algorithm 1, guarantees that

$$\|\hat{x}(k) - x(k)\|_2 \leq \kappa^c_{\mathsf{p},\mathsf{q},\mathsf{r}}(\Phi)\, v_{\max}(k), \quad {}^\forall k \geq 0$$

where $\lim_{k\to\infty} v_{\max}(k) = w_{\max}$.

Remark 3: For the resilient state estimation, most of the computational burden originates from the process of solving the optimization problem. The proposed decoder reduces the computational effort by combining the attack detection mechanism with the optimization process. Algorithm 1 only requires the minimization problem to be solved for a very short time interval when the attacker first attempts to inject false data so that the decoder has $f > \mathsf{q}$ at that instant. On the other

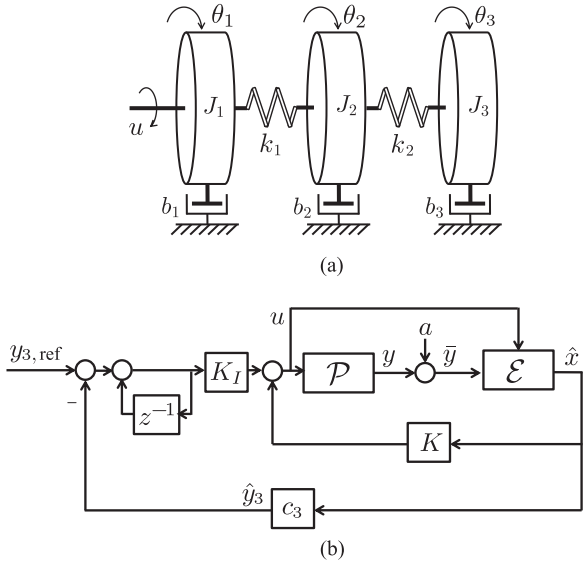

Fig. 3. Three inertia system and its control scheme. (a) System model. (b) Control block diagram.

hand, the estimator works as if there is no attack and computes only one simple pseudoinverse of a matrix during normal operation when $f \leq \mathsf{q}$ is guaranteed. ♦

Remark 4: Other observer-based resilient state estimators such as those in [4] and [14], consist of all possible combinations of estimator candidates. Thus, they need to run $\binom{\mathsf{p}}{\mathsf{q}}$ estimators so that the required memory size is $\mathsf{n}\binom{\mathsf{p}}{\mathsf{q}}$. On the other hand, the total memory size of all partial observers in the proposed estimator, $\sum_{\mathsf{i}=1}^{\mathsf{p}} \nu_{\mathsf{i}}$, is not greater than np because the size of each partial observer $\mathcal{O}_{\mathsf{i}}$ is $\nu_{\mathsf{i}} \leq \mathsf{n}$ for all $\mathsf{i} \in [\mathsf{p}]$. ♦

V. Simulation Results: Three Inertia System

To verify the effectiveness of the proposed scheme, simulations with a three inertia system are conducted in this section. The configuration of the three inertia system is described in Fig. 3(a) and its dynamics can be represented by a continuous-time state-space equation

$$\mathcal{P}_c : \begin{cases} \dot{x}(t) = A_c x(t) + B_c u(t) + d(t) \\ y(t) = C_c x(t) + n(t) + a(t) \end{cases} \tag{20}$$

with the matrices

$$A_c = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ -\frac{k_1}{J_1} & -\frac{b_1}{J_1} & \frac{k_1}{J_1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ \frac{k_1}{J_2} & 0 & -\frac{k_1+k_2}{J_2} & -\frac{b_2}{J_2} & \frac{k_2}{J_2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{k_2}{J_3} & 0 & -\frac{k_2}{J_3} & -\frac{b_3}{J_3} \end{bmatrix}$$

$$B_c = \begin{bmatrix} 0 \\ \frac{1}{J_1} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad C_c = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 \end{bmatrix}$$

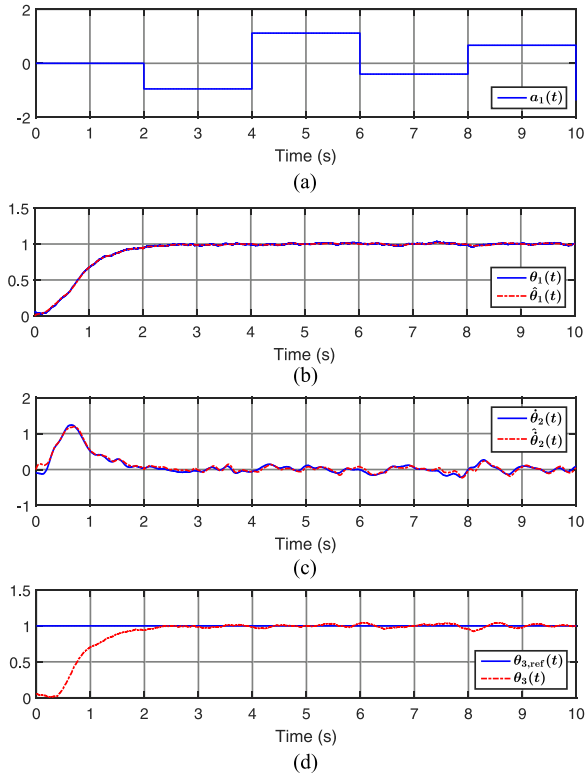

Fig. 4. Plot of signals. (a) $a_1(t)$. (b) $\theta_1(t)$ and $\hat{\theta}_1(t)$. (c) $\dot{\theta}_2(t)$ and $\hat{\dot{\theta}}_2(t)$. (d) $\theta_{3,\mathrm{ref}}(t)$ and $\theta_3(t)$.

where $J_1 = J_2 = J_3 = 0.01$ kg·m^2, $b_1 = b_2 = b_3 = 0.007$ N·m/(rad/s), and $k_1 = k_2 = 1.37$ N·m/rad. Here, the state variables are $x := [\theta_1 \ \dot{\theta}_1 \ \theta_2 \ \dot{\theta}_2 \ \theta_3 \ \dot{\theta}_3]^\top$ and the output measurements are $y := [\theta_1 \ \theta_2 \ \theta_3 \ \theta_1 - \theta_2 \ \theta_2 - \theta_3]^\top$. In addition, the plant is corrupted by the uniformly bounded process disturbance d and measurement noise n with $d_{\max} = n_{\max} = 0.001$. To conduct a discrete-time simulation, the zero-order hold equivalent model of (20) is considered, that is, the matrices of the discrete-time system (1) are given by $A := e^{A_c T_s}$, $B := \left(\int_0^{T_s} e^{A_c \tau} d\tau \right) B_c$, and $C := C_c$, where $T_s := 1$ ms denotes the sampling time. Note that the pair (A, C) is 2-redundant observable, which implies that one can correct the 1-sparse attack signal and its dynamic security index becomes 3. The control objective is to make the output θ_3 follow the step reference $\theta_{3,\mathrm{ref}}$. To this end, an observer-based feedback integral control scheme is adopted, as illustrated in [21, Section 6.7] and also in Fig. 3(b). First, the state feedback gains K and K_I are chosen as

$$K := -[2.32 \quad 0.25 \quad -2.47 \quad 0.04 \quad 1.70 \quad 0.12], \quad K_I := 0.002$$

as if the state x is available. Then, instead of using the conventional Luenberger observer, the proposed estimator $\mathcal{E}$ provides the estimate $\hat{x}$ of x. The injection gains L_{i} of the partial observer (14) in $\mathcal{E}$ are chosen arbitrarily such that $F_{\mathsf{i}} = S_{\mathsf{i}} - L_{\mathsf{i}} t_{\mathsf{i}}$ is Schur stable. Attack signals are illustrated in Fig. 4(a), which describes that adversaries launch a measurement data injection attack at $t = 2$ s so that the first sensor is compromised. Fig. 4(b) and (c) show state trajectories $\theta_1(t)$, $\dot{\theta}_2(t)$, and their estimates. It demonstrates the attack-resilient property of our estimation algorithm. Finally, Fig. 4(d) shows the reference tracking performance of the proposed control scheme.

VI. Conclusion

An LTI system is said to be $2\mathsf{q}$-redundant observable if it is observable even after eliminating any $2\mathsf{q}$ measurements. Relationships between the redundant observability and the security problems on cyber-physical systems under sensor attacks have been examined. To summarize, $2\mathsf{q}$-redundant observability implies that the numbers of detectable and correctable sensor attacks are $2\mathsf{q}$ and q, respectively. In addition, the dynamic security index, the minimum number of attacks to remain undetectable, is $2\mathsf{q}+1$.

Assuming that the measurement data injection attack is q-sparse and the disturbances/noises are bounded, an attack-resilient and robust state estimation scheme has been proposed under $2\mathsf{q}$-redundant observability. The proposed estimator consists of a bank of partial observers operating based on the Kalman observability decomposition and a decoder exploiting error correction techniques. In terms of time complexity, the decoder reduces the required computational effort by reducing the search space to a finite set and by combining a detection algorithm with the optimization process. On the other hand, in terms of space complexity, the required memory is linear with the number of sensors by means of the decomposition used for constructing a bank of partial observers.

References

[1] C. Lee, H. Shim, and Y. Eun, "Secure and robust state estimation under sensor attacks, measurement noises, and process disturbances: Observer-based combinatorial approach," in *Proc. 14th Eur. Control Conf.*, 2015, pp. 1866–1871.

[2] R. Langner, "Stuxnet: Dissecting a cyberwarfare weapon," *IEEE Security Privacy*, vol. 9, no. 3, pp. 49–51, May/Jun. 2011.

[3] Y. Mo *et al.*, "Cyber-physical security of a smart grid infrastructure," *Proc. IEEE*, vol. 100, no. 1, pp. 195–209, Jan. 2012.

[4] 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, Nov. 2013.

[5] H. Sandberg, S. Amin, and K. H. Johansson, "Cyberphysical security in networked control systems," *IEEE Control Syst.*, vol. 35, no. 1, pp. 20–23, Feb. 2015.

[6] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Trans. Inf. Syst. Security*, vol. 14, no. 1, pp. 13:1–13:33, May 2011.

[7] J. M. Hendrickx, K. H. Johansson, R. M. Jungers, H. Sandberg, and K. C. Sou, "Efficient computations of a security index for false data attacks in power networks," *IEEE Trans. Autom. Control*, vol. 59, no. 12, pp. 3194–3208, Dec. 2014.

[8] Y. Chen, S. Kar, and J. M. F. Moura, "Cyber-physical systems: Dynamic sensor attacks and strong observability," in *Proc. 40th IEEE Int. Conf. Acoust. Speech Signal Process.*, 2015, pp. 1752–1756.

[9] 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, Jun. 2014.

[10] M. Pajic *et al.*, "Robustness of attack-resilient state estimators," in *Proc. IEEE/ACM 5th Int. Conf. Cyber-Phys. Syst.*, 2014, pp. 163–174.

[11] 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, Aug. 2016.

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

[13] Y. Shoukry *et al.*, "SMT-based observer design for cyber-physical systems under sensor attacks," *ACM Trans. Cyber-Phys. Syst.*, vol. 2, no. 1, pp. 5:1–5:27, Feb. 2018.

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

[15] H. Jeon, S. Aum, H. Shim, and Y. Eun, "Resilient state estimation for control systems using multiple observers and median operation," *Math. Problems Eng.*, vol. 2016, Article no. 3750264, 2016.

[16] G. Basile and G. Marro, "On the observability of linear, time-invariant systems with unknown inputs," *J. Optim. Theory Appl.*, vol. 3, no. 6, pp. 410–415, Nov. 1969.

[17] Z. Gao, X. Liu, and M. Z. Q. Chen, "Unknown input observer-based robust fault estimation for systems corrupted by partially decoupled disturbances," *IEEE Trans. Ind. Electron.*, vol. 63, no. 4, pp. 2537–2547, Apr. 2016.

[18] C. Lee, H. Shim, and Y. Eun, "On redundant observability: From security index to attack detection and resilient state estimation," *ArXiv Preprint*, arXiv:1805.02640 [cs.SY], May 2018.

[19] V. Guruswami, J. R. Lee, and A. Wigderson, "Euclidean sections of ℓ_1^N with sublinear randomness and error-correction over the reals," in *Proc. 11th Int. Workshop APPROX 12th Int. Workshop RANDOM*, vol. 5171 (Lecture Notes in Computer Science), Berlin, Germany: Springer-Verlag, 2008, pp. 444–454.

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

[21] K. Ogata, *Discrete-Time Control Systems*, 2nd ed. Englewood Cliffs, NJ, USA: Prentice Hall, 1995.