State Estimation under Joint False Data Injection Attacks: Dealing with Constraints and Insecurity

Wenying Xu, *Member, IEEE,* Zidong Wang, *Fellow, IEEE,* Liang Hu, and Jürgen Kurths

***Abstract*—This paper is concerned with the security issue in the state estimation problem for a networked control system (NCS). A new model of joint false data injection (FDI) attack is established wherein attacks are injected to both the remote estimator and the communication channels. Such a model is general that includes most existing FDI attack models as special cases. The joint FDI attacks are subjected to limited access and/or resource constraints, and this gives rise to a few attack scenarios to be examined one by one. Our objective is to establish the so-called insecurity conditions under which there exists an attack sequence capable of driving the estimation bias to infinity while bypassing the anomaly detector. By resorting to the generalized inverse theory, necessary and sufficient conditions are derived for the insecurity under different attack scenarios. Subsequently, easy-to-implement algorithms are proposed to generate attack sequences on insecure NCSs with respect to different attack scenarios. In particular, by using a matrix splitting technique, the constraint-induced sparsity of the attack vectors is dedicatedly investigated. Finally, several numerical examples are presented to verify the effectiveness of the proposed FDI attacks.**

***Index Terms*—False data injection attack, security, joint attacks, state estimation, resource constraints.**

I. Introduction

For networked control systems (NCSs), the security is always a major concern as most communication channels, which form the backbones of NCSs, are vulnerable to cyber-attacks from malicious adversaries. As a matter of fact, a variety of security risks and threats have been found in industrial control systems including power grids [8], [16], [23], nuclear factories [19], and transportation systems [12], [13]. Typical cyber-attacks on NCSs include denial-of-service (DoS) attacks and deception attacks, where the *DoS attack* violates data *availability* by blocking data transmission among networked components [9], [28], [32], [36], [37], and the *deception attack* compromises information *integrity* through modifying data packets [18], [21], [23], [25], [39]. Compared with DoS attacks, deception attacks are more difficult to be detected since they are usually deliberately designed to bypass the anomaly detector. Two representative deception attacks are the replay attack [33], [35], [39] and the false data injection (FDI) attack [18], [20], [23], and the latter is the main focus of this paper.

State estimation problems for NCSs under deception attacks have been investigated mainly from two different but interrelated perspectives, namely, *the defenders* and *the attackers*. For typical defenders, they aim to develop estimation methods capable of withstanding certain types of deception attacks such as randomly occurring attacks [6], [7], [15], [34], unknown but bounded attacks [6], [24], and energy-constrained attacks [31], [39]. In particular, Bernoulli-distributed white sequences (with known conditional probabilities) have been proposed to characterize the success ratio of randomly occurring attacks in [15], [24].

In theory, a cunning attacker could purposely design the attack sequence that does not conform to any existing model, thereby resulting in divergent estimation errors. For instance, it has been found in [23] that the adversary can bypass the anomaly detector to deteriorate the estimation performance of power systems through deliberately launching a stealthy yet malicious attack. In this case, instead of looking into defenders' strategies of modeling malicious attacks under unavoidably strict assumptions, it might make more practical sense to examine the security problems from *the attacker*'s perspective, that is, assess the impact of deception attacks on the system performance without having to making possibly unrealistic assumptions. For some representative literature on attacker-based security problems, we refer the readers to [18], [29], [30], [38].

In terms of an attacker, the attention is focused on analyzing and assessing how different attacks affect the system performance. For *dynamic* NCSs, the impacts of attacks are typically dependent on where the attack is launched and the kinds of data compromised. Considering FDI attacks at *the communication channels* (where the measurement data is compromised), a computational method has been provided in [26] to quantify the maximal performance degradation on dynamic systems, and an insecurity condition has been derived in [18] for dynamic systems under FDI attacks. On the other hand, in consideration of the FDI attacks at *the estimator* (whose contents could be deleted, modified or corrupted), a necessary and sufficient condition has been proposed in [27] for the existence of undetectable attacks. It is worth noting that most existing results along this line have been based on the *single-FDI attack* scenario where only one component (either the communication channels or the estimator) of the NCS is compromised.

In engineering practice, a networked system consists of many components (e.g., controllers, estimators, sensors and actuators) connected via shared networks where the information exchanges are conducted via communication channels. Clearly, malicious FDI attacks could take place on any vulnerable component or communication channel. Such kind of attacks, referred to as *joint-FDI attacks*, compromises several (more than one) networked components/channels in a coordinated manner, thereby posing more serious threats to NCSs than their single-FDI counterparts. Nevertheless, despite its clear engineering insight, the security problem under joint-FDI attacks has so far been largely overlooked due probably to the mathematical complexities induced by the strong couplings of the joint attacks, and this constitutes one motivation of our current investigation.

Attacks, either on the communication channel or on the estimator, are inevitably subject to physical constraints. For example, some critical meters/sensors in power systems are protected against unauthorized physical access [22] and, in this case, only a subset of meters/sensors are approachable by adversaries. In practice, such

This work was supported in part by the Natural Science Foundation of Jiangsu Province under Grant BK20180367, the National Natural Science Foundation of China under Grant Nos. 62173087, 61933007, 61873148 and 61803082, the Fundamental Research Funds for the Central Universities of China, and in the Alexander von Humboldt Foundation of Germany. This work was also supported by ZhiShan Youth Scholar Program from Southeast University of China. *(Corresponding author: Zidong Wang)*.

W. Xu is with the School of Mathematics, Southeast University, Nanjing 211189 (e-mail: wenyingxuwinnie@gmail.com, wyxu@seu.edu.cn).

Z. Wang is with the Department of Computer Science, Brunel University London, Uxbridge, UB8 3PH, United Kingdom (email: Zidong.Wang@brunel.ac.uk).

L. Hu is with School of Computer Science and Electronic Engineering, University of Essex, Colchester, CO4 3SQ, United Kingdom (email: l.hu@essex.ac.uk).

J. Kurths is with Research Domain Complexity Science, Potsdam Institute for Climate Impact Research, 14412 Potsdam, Germany. He is also with Lobachevsky University of Nizhny Novgorod, Nizhnij Novgorod 603950, Russia (email: kurths@pik-potsdam.de).

physical constraints can be formulated as r-specific, that is, the attacks are only allowed to be launched on r-specific registers of the channels/estimator due probably to enhanced physical protection for others. These constraints can also be described in the form of r-sparsity, that is, at most r registers of the channels/estimator registers are vulnerable to injected attacks by the adversaries due to limited energy/resource/capacity [30]. Apparently, all these constraints would imply sparsity of the attack vector which, in turn, lead to substantial difficulties in the security research. In fact, together with the joint attack scenarios, the sparsity of the attack vector constitutes the major challenge in examining the attacks' impact on the NCSs.

Stimulated by the discussions made so far, in this paper, we aim to investigate the impact from a variety of constrained joint-FDI attacks on the estimation performance of a class of NCSs. Insecurity conditions, under which there exist undetectable attack sequences resulting in unbounded estimation bias, are established by virtue of generalized inverse theory. The contribution of this paper can be highlighted as fourfold: *1) a rather general joint-FDI attack model is proposed, for the first time, to characterize the settings where false data is injected into both the communication channels and the estimator of the NCS; 2) both the limited access and constrained resource capacity are taken into special account for the adversary in order to better reflect the reality of the FDI attacks; 3) necessary and sufficient conditions for the insecurity are derived for the state estimation problem of NCSs under various FDI attacks; and 4) easily implementable algorithms are put forward, by resorting to matrix splitting techniques, to design attack sequences under different attack scenarios with specific efforts to tackle the sparsity in constrained attacks.*

Notations: $\mathbb{R}^n$ and $\mathbb{R}^{n\times m}$ stand for the n-dimensional Euclidean space and $n\times m$ real matrices, respectively. For a matrix M, $\|M\|$, $\mathrm{Image}(M)$ and $\mathrm{Ker}(M)$ mean the Frobenius norm, the image space and the kernel space of M, respectively, and $M\succ 0$ ($\succeq 0$) means that M is positive definite (semi-definite). For a set $S\subseteq\mathbb{R}^n$, the complement of S is denoted as $\overline{S}\triangleq\{x\in\mathbb{R}^n|x\notin S\}$. Moreover, for $\forall x\in\mathbb{R}^n$, $\|x\|$ denotes its l_2 norm, and x is called a k-sparse vector if it has at most k nonzero elements. Let $\{x(k)\}$ denote an infinite sequence of $x(1)$, $x(2)$, $\cdots$, $x(k)$, $\cdots$. Furthermore, I and $\mathbb{O}$ represent, respectively, the identity matrix and zero matrix with compatible dimension. $\mathbf{1}$ (or $\mathbf{0}$) is a vector with all elements being 1 (or 0), and $\mathbf{1}_m^s=[\overbrace{0,\cdots,0}^{s-1},\underbrace{1,0,\cdots,0}_{m}]^T$.

II. Problem Formulation and Preliminaries

In this section, we present some preliminaries related to FDI attacks and NCSs, and then introduce the problem setup.

A. State estimation without attacks

Consider a discrete stochastic linear time-invariant (LTI) physical plant of the following form:

$$\left\{\begin{array}{rcl} x(k+1) &=& Ax(k)+w(k)\\ y(k) &=& Cx(k)+v(k)\end{array}\right. \tag{1}$$

where $x(k)\in\mathbb{R}^n$ and $y(k)=[y_1(k),\cdots,y_m(k)]^T\in\mathbb{R}^m$ are the system state and the measurement output at time k, respectively; $y_i(k)$ is the measurement of sensor i at time k transmitted over channel i ($i=1,2,\cdots,m$) (see Fig. 1); the process noise $w(k)\in\mathbb{R}^n$ and the measurement noise $v(k)\in\mathbb{R}^m$ are mutually uncorrelated zero-mean random signals with covariance matrices $Q\succeq 0$ and $R\succ 0$, respectively; the initial state $x(0)$ is a zero-mean Gaussian random variable with covariance $\Sigma\succeq 0$, which is independent of $w(k)$ and $v(k)$; and the system matrices A and C are known with the appropriate dimensions. The pair (C,A) is observable and $(A,Q^{\frac{1}{2}})$ is controllable.

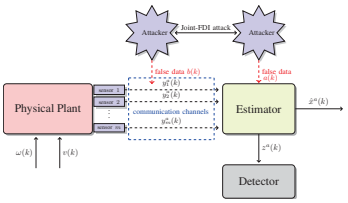

Fig. 1. Diagram of state estimation under joint-FDI attacks.

A Kalman filter is introduced to compute the state estimate $\hat{x}(k)$ from the measurement output $y(k)$, i.e.,

$$\hat{x}(k+1|k)=A\hat{x}(k)\quad P_{k+1|k}=AP_kA^T+Q \tag{2}$$

$$K_k=P_{k|k-1}C^T(CP_{k|k-1}C^T+R)^{-1} \tag{3}$$

$$\hat{x}(k)=\hat{x}(k|k-1)+K_k(y(k)-C\hat{x}(k|k-1)) \tag{4}$$

$$P_k=P_{k|k-1}-K_kCP_{k|k-1} \tag{5}$$

with the initial state $\hat{x}(0)$ and the initial value P_0. It is well known that the estimation error covariance of the Kalman filter will converge to a value if (C,A) is observable and $(A,Q^{\frac{1}{2}})$ is controllable [3]. In this paper, we assume that the Kalman filter gain K is operating at a steady state with $\triangleq\lim_{k\to\infty}P_{k|k-1}$ and $K\triangleq PC^T(CPC^T+R)^{-1}$. Then, $A-KCA$ is stable [3].

From (2) and (4), it follows that the recursion of $\hat{x}(k)$ satisfies

$$\hat{x}(k+1)=A\hat{x}(k)+K(y(k+1)-CA\hat{x}(k)) \tag{6}$$

and the estimation residual $z(k+1)$ at time $k+1$ is given by

$$z(k+1)=y(k+1)-CA\hat{x}(k). \tag{7}$$

Letting $e(k)=x(k)-\hat{x}(k)$ be the estimation error, we have

$$e(k+1)=(A-KCA)e(k)+(I_n-KC)w(k)-Kv(k+1). \tag{8}$$

Note that system (8) is stable if and only if $A-KCA$ is stable [17]. For the rest of the discussion, the system (8) is assumed to be in a steady state.

In this paper, a χ^2 false-data detector is utilized for the estimator to diagnose the potential existence of attacks. At time k, the χ^2 false-data detector checks whether $g(k)>\alpha_0$ holds or not with a prescribed threshold α_0 and $g(k)=z^T(k)(CPC^T+R)^{-1}z(k)$. If $g(k)>\alpha_0$, then an alarm will be triggered, which implies that the system is under attack, otherwise the system is thought to operate normally.

B. Joint-FDI attack model

In this subsection, we describe a joint-FDI attack model with false data injected into both the communication channels and the estimator (see Fig. 1). Before introducing the attack model, the following assumptions are made on the malicious adversary.

Assumption 1. *The adversary knows the system parameters, i.e., the matrices A and C.*

Assumption 2. *The adversary is able to inject a false data sequence $\{a(k)\}$ into the estimator register, that is, at time k, the state estimate $\hat{x}(k-1)$ can be reset to $\hat{x}(k-1)+a(k)$ by the adversary.*

Assumption 2, adopted from [27], assumes that the adversary is capable of modifying or corrupting the contents stored in the register of the estimator.

Assumption 3. *The adversary is able to inject a false data sequence $\{b(k)\}$ into the communication channels between the sensors and the estimator.*

1) Joint-FDI attack: The false data $a(k)$ and $b(k)$ are injected into the estimator and communication channels at time k, respectively. Under such an attack, the measurement output received by the estimator is given by

$$y^a(k) = Cx(k) + b(k) + v(k) \tag{9}$$

and the dynamics of the estimator is given as follows:

$$\begin{cases} \hat{x}^a(k+1) = A(\hat{x}^a(k) + a(k+1)) + Kz^a(k+1) \\ z^a(k+1) = y^a(k+1) - C[A(\hat{x}^a(k) + a(k+1))] \end{cases} \tag{10}$$

where $y^a(k) \in \mathbb{R}^m$, $\hat{x}^a(k) \in \mathbb{R}^n$ and $z^a(k) \in \mathbb{R}^m$ are, respectively, the measurement output, the state estimate and the estimation residual of system (1) under attack at time k. The vectors $a(k) = [a_1(k), \cdots, a_n(k)]^T \in \mathbb{R}^n$ and $b(k) = [b_1(k), \cdots, b_m(k)]^T \in \mathbb{R}^m$ represent the false data injected into the estimator and the channels at time k, respectively.

Remark 1. *In real-world NCSs, it is quite common that several (more than one) networked components are subject to coordinated attacks as a result of the advances in intrusion techniques. In* (9)-(10)*, we adopt two attack vectors $a(k)$ and $b(k)$ to characterize joint-FDI attacks. In cases of $a(k) = 0$ or $b(k) = 0$, the joint-FDI attacks reduce to the single-FDI ones that have already been discussed in [18], [26], [27] (see Table I). As such, our joint-FDI model is more reasonable (yet more general) than those existing single-FDI ones.*

Next, by letting $b(k) \equiv \mathbf{0}$ in joint-FDI attacks, we present a kind of single-FDI attacks that will be addressed subsequently.

2) Single-FDI attack: The bad data $a(k)$ is injected into the estimator at time k. Under such an attack, the dynamics of the estimator (10) becomes

$$\begin{cases} \hat{x}^a(k+1) = A\left(\hat{x}^a(k) + a(k+1)\right) + Kz^a(k+1) \\ z^a(k+1) = y(k+1) - C[A(\hat{x}^a(k) + a(k+1))]. \end{cases} \tag{11}$$

In the above model, all communication channels are protected and the received measurement output is $y(k+1)$ (rather than $y^a(k)$ in the case of joint-FDI attack).

C. Attack scenarios

In this paper, we consider the case where cyber-attacks can be launched on both the remote estimator and the communication channels, and the attacks themselves are subjected to limited access and/or resource constraints. In terms of the places where the attacks take place and the access/resource constraints, we have six different attacks (see Table I) that cover a fairly wide range of scenarios. For each scenario $\mathcal{S}_i$ $(i = 1, 2, \cdots, 6)$, we formulate the corresponding attack strategy $J(\mathcal{S}_i)$ that presents a feasible set for all possible attack vectors $a(k)$ and $b(k)$. To be specific, the six attack scenarios include two *ideal* ones $\mathcal{S}_i$ $(i = 1, 2)$ and four *constrained* scenarios $\mathcal{S}_i$ $(i = 3, 4, 5, 6)$ (see Table I). The four *constrained* scenarios can be described as follows:

- $\mathcal{S}_3$: the joint-FDI attack is launched with full access to the estimator but limited access to r specific communication channels;
- $\mathcal{S}_4$: the joint-FDI attack is launched with full access to the estimator but limited resources to compromise up to r communication channels (out of all the channels), i.e., $b(k)$ is restricted to a $r-$sparse vector;
- $\mathcal{S}_5$: the single-FDI attack is launched with limited resources to compromise a subset of the estimator registers, i.e., $a(k)$ is restricted to a $r-$sparse vector;
- $\mathcal{S}_6$: the joint-FDI attack is launched with limited resources to compromise a subset of the estimator registers and the communication channels, i.e., $a(k)$ and $b(k)$ are restricted to $r-$sparse vectors.

Remark 2. *In reality, malicious attacks are likely to undergo ineluctable constraints resulting from a number of factors such as physical protection of the targeted system [22], limited energy supply for attacks and resource capacity for adversaries [30]. For this reason, the so-called $r-$specific vectors $a(k)$ and $b(k)$ (with r specific elements being 0) are introduced as a true reflection of limited physical access induced by system protection. Also, $a(k)$ and $b(k)$ are restricted to $r-$sparse vectors with aim to characterize the impact of constrained resource capacity on attacks.*

Define the state estimation difference $\Delta x(k)$ and the estimation residual difference $\Delta z(k)$ as follows:

$$\begin{aligned} \Delta x(k+1) &\triangleq \hat{x}^a(k+1) - \hat{x}(k+1) \\ \Delta z(k+1) &\triangleq z^a(k+1) - z(k+1). \end{aligned} \tag{12}$$

Then, we provide the definition of an insecure system.

Definition 1. *[18] The system* (1) *with the state estimator* (6) *is said to be insecure in the attack scenario $\mathcal{S}_i$ $(i = 1, \cdots, 6)$ if there exist attack sequences $\{a(k), b(k)\} \subseteq J(\mathcal{S}_i)$ such that the following two conditions hold simultaneously:*

(1) *The state estimation difference $\Delta x(k)$ satisfies*

$$\lim_{k\to\infty} \|\Delta x(k)\| = \infty, \tag{13}$$

(2) *The estimation residual difference $\Delta z(k)$ satisfies*

$$\|\Delta z(k)\| \leq \alpha \tag{14}$$

where α is a prescribed small positive constant scalar.

It is worthwhile to mention that the definition of system security (i.e. Definition 1) is the same as that of system vulnerability (see [29, Definition 2]). From Definition 1 and [29, Definition 2]), a system is secure (or invulnerable) if it is always stable under any attacks that have bounded influence on the residue, that is, $\|\Delta z(k)\| \leq \alpha$.

This paper aims to investigate the impact from a variety of constrained joint-FDI attacks on the estimation performance of a class of NCSs. To the best of our knowledge, this problem remains challenging with the following two substantial difficulties to be resolved: i) the establishment of necessary and sufficient conditions (occasionally sufficient conditions only) for the existence of attack vectors, which are capable of driving the bias in state estimate to infinity but bypassing the anomaly detector; and ii) the design of attack sequences under different attack scenarios with specific efforts to tackle the sparsity in constrained attacks. Therefore, our main tasks are to deal with these two emerging difficulties.

III. Ideal Attack Scenarios

In this section, the ideal attack scenarios $\mathcal{S}_1$ for the single-FDI attack and $\mathcal{S}_2$ for the joint-FDI attack are investigated, respectively.

For the convenience of theoretical analysis, we provide the following lemmas.

Lemma 1. *[10] Let V and W be finite dimensional vector spaces, and $T : V \to W$ is a linear transformation. Then, we have*

$$\mathrm{Rank}(T) + \mathrm{Nullity}(T) = \dim V. \tag{15}$$

Moreover, if T is idempotent, i.e., $T^2 = T$, then $\mathrm{Ker}(T) = \mathrm{Image}(I - T)$ *where* $\mathrm{Rank}(T) \triangleq \dim(\mathrm{Image}(T))$, $\mathrm{Nullity}(T) \triangleq \dim(\mathrm{Ker}(T))$. *Here, the kernel space and the image space are*

SUMMARY OF ATTACK SCENARIOS

Scenario	Attack mode	Channels	Estimator	
Ideal $\mathcal{S}_1$	single-FDI	$\times$	$\surd$	[27], Proposition 1, Corollary 1
Ideal $\mathcal{S}_1'$	single-FDI	$\surd$	$\times$	[18], [26]
Ideal $\mathcal{S}_2$	joint-FDI	$\surd$	$\surd$	Theorem 1
Constrained $\mathcal{S}_3$	joint-FDI	r-specific	$\surd$	Theorem 2
Constrained $\mathcal{S}_4$	joint-FDI	r-sparse	$\surd$	Theorem 3
Constrained $\mathcal{S}_5$	single-FDI	$\times$	r-sparse	Theorem 4
Constrained $\mathcal{S}_6$	joint-FDI	r-sparse	r-sparse	Theorem 5

$\surd$ full access $\times$ no access

defined as $\mathrm{Ker}(T) = \{v \in V \ : \ Tv = \mathbf{0}\}$ *and* $\mathrm{Image}(T) = \{Tv \ : \ v \in V\}$*, respectively.*

Lemma 2. *[10] For matrices $A \in \mathbb{R}^{n\times m}$ and $B \in \mathbb{R}^{m\times r}$, the following inequality holds:* $\mathrm{Rank}(AB) \leq \min\{\mathrm{Rank}(A), \mathrm{Rank}(B)\}$.

Definition 2. *[10] Given a matrix $A \in \mathbb{R}^{n\times m}$, a matrix $A^g \in \mathbb{R}^{m\times n}$ is said to be a generalized inverse of A if $AA^gA = A$.*

Lemma 3. *There exists a matrix T such that $TCA = A$ if and only if* $\mathrm{Rank}(CA) = \mathrm{Rank}(A)$. *Here $T = A(CA)^g$ and $(CA)^g$ is the generalized inverse of CA.*

Proof: (Sufficiency) By the definition of the generalized inverse, one has $\mathrm{Rank}(CA) = \mathrm{Rank}(CA(CA)^gCA) \leq \mathrm{Rank}((CA)^gCA) \leq \mathrm{Rank}(CA)$. Note that $\mathrm{Rank}(CA) = \mathrm{Rank}(A)$, then $\mathrm{Rank}((CA)^gCA) = \mathrm{Rank}(A)$. It follows from Lemma 1 that $\mathrm{Nullity}(A) = \mathrm{Nullity}((CA)^gCA)$. Then, one concludes that $\mathrm{Ker}(A) = \mathrm{Ker}((CA)^gCA)$ due to $\mathrm{Ker}(A) \subseteq \mathrm{Ker}((CA)^gCA)$. Let $T \triangleq (CA)^gCA$. Obviously, T is idempotent since $T^2 = T$. It follows from Lemma 1 that $\mathrm{Ker}(A) = \mathrm{Ker}(T) = \mathrm{Image}(I-T)$, which implies that $(I-T)v \in \mathrm{Ker}(A)$ for $\forall v \in \mathbb{R}^n$. Because v is arbitrary, we conclude that $A(I-T) = \mathbb{O}$, namely, $A(CA)^gCA = A$.

(Necessity) The necessity follows immediately from $\mathrm{Rank}(A) \leq \mathrm{Rank}(CA) \leq \mathrm{Rank}(A)$. ■

A. Single-FDI attack in scenario $\mathcal{S}_1$

In the ideal attack scenario $\mathcal{S}_1$, the attack has full access to the estimator only. The corresponding attack strategy $J(\mathcal{S}_1)$ is given by

$$J(\mathcal{S}_1) = \{\{a(k), b(k)\} \mid a(k) \in \mathbb{R}^n, b(k) \equiv \mathbf{0}, k \in \mathbb{N}^+\}. \quad (16)$$

The dynamics of $\Delta x(k+1)$ and $\Delta z(k+1)$ can be obtained from (11)-(12) as follows:

$$\begin{array}{ll} \Delta x(k+1) & = (I-KC)A\left[\Delta x(k) + a(k+1)\right] \\ \Delta z(k+1) & = -CA(\Delta x(k) + a(k+1)). \end{array} \quad (17)$$

Next, we discuss the necessary and sufficient condition under which the system (1) with the state estimator (6) is insecure in $\mathcal{S}_1$.

Proposition 1. *The system* (1) *with the state estimator* (6) *is insecure in the attack scenario $\mathcal{S}_1$ if and only if* $\mathrm{Rank}(CA) < \mathrm{Rank}(A)$.

Proof: According to Lemma 1, $\mathrm{Rank}(CA) < \mathrm{Rank}(A)$ if and only if $\mathrm{Nullity}(CA) > \mathrm{Nullity}(A)$. It is worthwhile to mention that $CA\eta = \mathbf{0}$ if $A\eta = \mathbf{0}$ for $\forall \eta \in \mathbb{R}^n$, which implies $\mathrm{Ker}(A) \subseteq \mathrm{Ker}(CA)$.

(Sufficiency) It follows from $\mathrm{Rank}(CA) < \mathrm{Rank}(A)$ that $\mathrm{Ker}(A) \subsetneq \mathrm{Ker}(CA)$. That is, there exists a vector $\eta^* \in \mathbb{R}^n$ that satisfies $\eta^* \in \mathrm{Ker}(CA)$ but $\eta^* \notin \mathrm{Ker}(A)$.

Construct the attack sequence $\{a(k)\}$ in the following form

$$a(k) = \beta_k\eta^* - \beta_{k-1}A\eta^* \quad (k \geq 2) \quad (18)$$

with $a(1) = \beta_1\eta^*$ and $\lim_{k\to\infty}\beta_k = \infty$. Notice that $\Delta x(0) = \mathbf{0}$. Substituting the attack sequence $\{a(k)\}$ in (18) into (17), it is not difficult to see $\|\Delta z(k)\| = 0$ but $\|\Delta x(k)\| = \beta_k\|A\eta^*\| \to \infty$ as $k \to \infty$ due to $A\eta^* \neq \mathbf{0}$. Thus, we conclude that the system (1) with the state estimator (6) is insecure in the scenario $\mathcal{S}_1$ according to Definition 1.

(Necessity) Let us prove the necessity by contradiction. Suppose that $\mathrm{Rank}(CA) \geq \mathrm{Rank}(A)$, then it follows from Lemma 2 that $\mathrm{Rank}(CA) = \mathrm{Rank}(A)$. According to Lemma 3, for $\forall \eta \in \mathbb{R}^n$, there exists a matrix $T \in \mathbb{R}^{n\times m}$ such that $A\eta = TCA\eta$. From (17), we have $\Delta x(k+1) = (I-KC)T\Delta z(k+1)$, which leads to $\|\Delta x(k+1)\| \leq \|(I-KC)T\|\|\Delta z(k+1)\|$. Then, $\|\Delta x(k+1)\|$ is bounded if $\|\Delta z(k+1)\|$ is bounded, and this violates the condition for the insecurity as stated in Definition 1. The proof is complete. ■

Corollary 1. *If* $\mathrm{Rank}(C) < \mathrm{Rank}(A)$*, then system* (1) *with state estimator* (6) *is insecure under the single-FDI attack in $\mathcal{S}_1$.*

Proof: According to Lemma 2, one has $\mathrm{Rank}(CA) < \mathrm{Rank}(A)$, and the proof follows readily from Proposition 1. ■

Note that $\mathrm{Rank}(C) < \mathrm{Rank}(A)$ is a sufficient (but not necessary) condition.

Remark 3. *Although the case of the single-FDI attack has already been discussed in [27], a new easy-to-check necessary and sufficient condition is derived in Proposition 1. Different from [27], more general FDI attacks will be discussed subsequently.*

B. Joint-FDI attack in scenario $\mathcal{S}_2$

A seemingly natural question arises as follows: if the conditions in Proposition 1 and Corollary 1 do not hold, is it possible for the adversary to make the system (1) (with the estimator (6)) insecure? To answer this question, a joint-FID attack is constructed with $a(k)$ and $b(k)$ injected into the estimator and the communication channels simultaneously. In this subsection, we consider the ideal attack scenario $\mathcal{S}_2$, where both the estimator and the channels are fully accessible. The corresponding attack strategy $J(\mathcal{S}_2)$ is given by $J(\mathcal{S}_2) = \{\{a(k), b(k)\} \mid a(k) \in \mathbb{R}^n, b(k) \in \mathbb{R}^m, k \in \mathbb{N}^+\}$. Obviously, $a(k)$ and $b(k)$ are arbitrary with given dimensions.

Taking (10) and (12) into consideration, we obtain

$$\begin{aligned} \Delta x(k+1) =& (I-KC)A\left[\Delta x(k) + a(k+1)\right] + Kb(k+1) \\ \Delta z(k+1) =& -CA(\Delta x(k) + a(k+1)) + b(k+1). \end{aligned} \quad (19)$$

Next, let us explore the necessary and sufficient condition under which the system (1) (with the state estimator (6)) is insecure in $\mathcal{S}_2$.

Theorem 1. *The system* (1) *with the state estimator* (6) *is insecure in the attack scenario $\mathcal{S}_2$ if and only if $A \neq \mathbb{O}$.*

Proof: (Sufficiency) If $A \neq \mathbb{O}$, then there exists a nonzero vector $\eta \in \mathbb{R}^n$ such that $A\eta \neq \mathbf{0}$. The construction procedure of the attack sequences $\{a(k)\}$ and $\{b(k)\}$ is presented in Algorithm 1, i.e.,

$$b(k) = \beta_k CA\eta \quad \text{and} \quad a(k+1) = \beta_{k+1}\eta - \beta_k A\eta \tag{20}$$

where, for $k \in \mathbb{N}^+$, $a(1) = \beta_1\eta$ and $\lim_{k\to\infty} \beta_k = \infty$. By noticing that $\Delta x(0) = \mathbf{0}$, one has $\Delta x(k) = \beta_k A$ and $\Delta z(k) = \mathbf{0}$. Obviously, $\|z(k)\| = 0 \leq \alpha$, but $\|\beta_k A\eta\| = \beta_k \|A\eta\| \to \infty$ as $k \to \infty$ due to $A\eta \neq \mathbf{0}$. Therefore, the conditions (13)-(14) are satisfied, and we conclude that the system (1) is insecure under the attacks generated by Algorithm 1.

Algorithm 1 Algorithm for the ideal scenario $\mathcal{S}_2$

Initialization: Set $a(1) = \beta_1\eta$

1: **while** $k \geq 1$ **do**
2: Set attack $b(k) = \beta_k CA\eta$
3: Set attack $a(k+1) = \beta_{k+1}\eta - \beta_k A\eta$
4: $k \leftarrow k+1$

(Necessity) We prove the necessity by contradiction. Suppose that $A = \mathbb{O}$. Then, (19) is equivalent to $\Delta x(k+1) = Kb(k+1)$ and $\Delta z(k+1) = b(k+1)$, which implies that $\|\Delta x(k+1)\| = \|Kz(k+1)\| \leq \|K\|\|z(k+1)\|$. As such, it is impossible for (13) and (14) to be true at the same time. The proof is complete. ∎

Remark 4. *A necessary and sufficient condition is established in Theorem 1, which guarantees the existence of attack sequences that are capable of driving the estimation bias to infinity but bypassing the anomaly detector. Specifically, if $A \neq \mathbb{O}$, then the system* (1) *with the estimator* (6) *is insecure, and an implementable procedure is presented in Algorithm 1 to generate a proper attack sequence. On the other hand, if the system* (1) *with the estimator* (6) *is insecure, then $A \neq \mathbb{O}$ which, in turn, reflects that the system* (1) *with $A = \mathbb{O}$ is robust to attacks launched in the scenario $\mathcal{S}_2$.*

Remark 5. *It is worthwhile to mention that, compared with single-FDI attacks discussed in Proposition 1, the joint-FDI attacks pose more serious threats to the system security due to a relaxed insecure condition derived in Theorem 1. Specifically, if the system* (1) *with the estimator* (6) *is insecure under the single-FDI attack, then $\mathrm{Rank}(CA) < \mathrm{Rank}(A)$ by Proposition 1, which implies that $A \neq \mathbb{O}$ and thus the system* (1) *is also insecure under the joint-FDI attack. On the other hand, the system* (1) *with the estimator* (6)*, which is insecure under the joint-FDI attack, is not necessarily insecure under the single-FDI attack, since $A \neq \mathbb{O}$ does not always imply $\mathrm{Rank}(CA) < \mathrm{Rank}(A)$.*

IV. Joint-FDI Attack With Constraints On $b(k)$

Two constrained attack scenarios $\mathcal{S}_3$ and $\mathcal{S}_4$ for joint-FDI attacks are discussed in this section.

For the convenience of analysis, we define two index subsets

$$\mathcal{I}_n^r = \{i_1, i_2, \cdots, i_r\}, \quad \mathcal{J}_n^r = \{j_1, j_2, \cdots, j_{n-r}\}$$

of the set $\{1, 2, \cdots, n\}$ satisfying

$$\mathcal{I}_n^r \cap \mathcal{J}_n^r = \varnothing, \quad \mathcal{I}_n^r \cup \mathcal{J}_n^r = \{1, 2, \cdots, n\}. \tag{21}$$

Furthermore, we define two matrices

$$E_{\mathcal{I}_n^r} = [\mathbf{1}_n^{i_1}, \mathbf{1}_n^{i_2}, \cdots, \mathbf{1}_n^{i_r}] \in \mathbb{R}^{n\times r} \tag{22}$$

$$E_{\mathcal{J}_n^r} = [\mathbf{1}_n^{j_1}, \mathbf{1}_n^{j_2}, \cdots, \mathbf{1}_n^{j_{n-r}}] \in \mathbb{R}^{n\times(n-r)}, \tag{23}$$

which have the following properties:

$$E_{\mathcal{I}_n^r}^T E_{\mathcal{I}_n^r} = I \quad E_{\mathcal{J}_n^r}^T E_{\mathcal{J}_n^r} = I \quad E_{\mathcal{I}_n^r}^T E_{\mathcal{J}_n^r} = \mathbb{O} \tag{24}$$

$$E_{\mathcal{J}_n^r}^T E_{\mathcal{I}_n^r} = \mathbb{O} \quad E_{\mathcal{I}_n^r} E_{\mathcal{I}_n^r}^T + E_{\mathcal{J}_n^r} E_{\mathcal{J}_n^r}^T = I. \tag{25}$$

A. Constrained scenario $\mathcal{S}_3$: limited access to specific communication channels

There are m communication channels between the sensors and the estimator. In this subsection, the scenario $\mathcal{S}_3$ is first discussed where m-r (out of m) channels are protected, and only r specific channels are accessible to the adversary. Without loss of generality, let $\mathcal{I}_{acc} = \{s_1^*, s_2^*, \cdots, s_r^*\}$ be the set of indices of the r channels approachable by the attackers, and let $\mathcal{I}_{pro} = \{s_1, s_2, \cdots, s_{m-r}\}$ be the set of indices of the protected channels.

The attack strategy $J(\mathcal{S}_3)$ with respect to the scenario $\mathcal{S}_3$ is given by $J(\mathcal{S}_3) = \big\{\{a(k), b(k)\} \mid a(k) \in \mathbb{R}^n, b_i = 0 \text{ with } i \in \mathcal{I}_{pro}, b(k) \in \mathbb{R}^m, k \in \mathbb{N}^+\big\}$. Under this strategy, the output measurement $y_i(k)$ ($i \in \mathcal{I}_{acc}$) is possibly manipulated to $y_i(k) + b_i(k)$ at instant k, while $y_j(k)$ ($j \in \mathcal{I}_{pro}$) can be perfectly transmitted over channel j.

Note that $\mathcal{I}_{acc} \cap \mathcal{I}_{pro} = \varnothing$ and $\mathcal{I}_{acc} \cup \mathcal{I}_{pro} = \{1, 2, \cdots, m\}$. Then, define two matrices $E_{acc} = [\mathbf{1}_m^{s_1^*}, \mathbf{1}_m^{s_2^*}, \cdots, \mathbf{1}_m^{s_r^*}] \in \mathbb{R}^{m\times r}$ and $E_{pro} = [\mathbf{1}_m^{s_1}, \mathbf{1}_m^{s_2}, \cdots, \mathbf{1}_m^{s_{m-r}}] \in \mathbb{R}^{m\times(m-r)}$ with $s_i^* \in \mathcal{I}_{acc}$ and $s_i \in \mathcal{I}_{pro}$.

Theorem 2. *The system* (1) *with the state estimator* (6) *is insecure in the attack scenario $\mathcal{S}_3$ if and only if* $\mathrm{Rank}(E_{pro}^T CA) < \mathrm{Rank}(A)$.

Proof: The proof is similar to that of Theorem 3 and is thus omitted here. ∎

Remark 6. *A necessary and sufficient condition is derived in Theorem 2 that reflects what system is insecure in the attack scenario $\mathcal{S}_3$. In addition, if the system* (1) *with the estimator* (6) *is insecure, then $\mathrm{Rank}(E_{pro}^T CA) < \mathrm{Rank}(A)$ which implies that, if the condition $\mathrm{Rank}(E_{pro}^T CA) = \mathrm{Rank}(A)$ is satisfied, then the system* (1) *with the estimator* (6) *is robust to the so-called access-constrained joint-FDI attacks launched in $\mathcal{S}_3$. Such a condition could help the system designer to understand what channels are critical to the system security and then do the needful to schedule the protection priority.*

Remark 7. *According to Theorem 2, in case of $E_{pro}^T = I$ (or $E_{pro}^T = \mathbb{O}$), the so-called access-constrained joint-FDI attack becomes the full-access single-FDI one (or joint-FDI one) that has been discussed in Proposition 1 (or Theorem 1). As such, Proposition 1 and Theorem 1 can be regarded as special cases of Theorem 2.*

B. Constrained scenario $\mathcal{S}_4$: limited available resources for compromising communication channels

In this subsection, we further consider the scenario $\mathcal{S}_4$, where the attack has the limited resource capacity for compromising up to r channels (out of all the channels). Note that here is no restriction on what channels are protected in $\mathcal{S}_4$ (as opposed to the case of $\mathcal{S}_3$). In other words, any r channels (out of m channels) are accessible to the adversary.

The attack strategy $J(\mathcal{S}_4)$ with respect to the scenario $\mathcal{S}_4$ is given by $J(\mathcal{S}_4) = \big\{\{a(k), b(k)\} \mid b(k) \in \mathbb{R}^m \text{ is a } r\text{-sparse vector}, a(k) \in \mathbb{R}^n, k \in \mathbb{N}^+\big\}$.

Theorem 3. *The system* (1) *with the state estimator* (6) *is insecure in the attack scenario $\mathcal{S}_4$ if and only if there exist two index sets $\mathcal{I}_m^r, \mathcal{J}_m^r$ with properties in* (21) *such that* $\mathrm{Rank}(E_{\mathcal{J}_m^r}^T CA) < \mathrm{Rank}(A)$.

Proof: According to Lemma 1, $\mathrm{Rank}(E_{\mathcal{J}_m^r}^T CA) < \mathrm{Rank}(A)$ if and only if $\mathrm{Nullity}(E_{\mathcal{J}_m^r}^T CA) > \mathrm{Nullity}(A)$. Note that $E_{\mathcal{J}_m^r}^T CA\xi =$

0 if $A\xi = \mathbf{0}$ for $\xi \in \mathbb{R}^n$, which implies that $\mathrm{Ker}(A) \subseteq \mathrm{Ker}(E_{\mathcal{J}_m^r}^T CA)$.

(Sufficiency) If $\mathrm{Rank}(E_{\mathcal{J}_m^r}^T CA) < \mathrm{Rank}(A)$, then $\mathrm{Ker}(A) \subsetneq \mathrm{Ker}(E_{\mathcal{J}_m^r}^T CA)$. That is, there exists a vector $\xi^* \in \mathbb{R}^n$ that satisfies $\xi^* \in \mathrm{Ker}(E_{\mathcal{J}_m^r}^T CA)$ but $\xi^* \notin \mathrm{Ker}(A)$, namely, $E_{\mathcal{J}_m^r}^T CA\xi^* = \mathbf{0}$ but $A\xi^* \neq \mathbf{0}$.

The construction procedure of $\{a(k), b(k)\}$ is displayed in Algorithm 2 with $\lim_{k\to\infty} \beta_k = \infty$. Next, we will illustrate that (i) $b(k)$ is a r-sparse vector, and (ii) the system (1) is insecure under the attack sequences $\{a(k), b(k)\}$.

First, $b(k)$ is a r-sparse vector due to $E_{\mathcal{I}_m^r}^T CA\xi^* \in \mathbb{R}^r$. Second, noting that $\Delta x(0) = \mathbf{0}$, it follows from Algorithm 2 that $\Delta x(k+1) = \beta_{k+1} A\xi^*$ and $\Delta z(k+1) = -\beta_{k+1} CA\xi^* + \beta_{k+1} E_{\mathcal{I}_m^r} E_{\mathcal{I}_m^r}^T CA\xi^* = -\beta_{k+1}(E_{\mathcal{I}_m^r} E_{\mathcal{I}_m^r}^T + E_{\mathcal{J}_m^r} E_{\mathcal{J}_m^r}^T)CA\xi^* + \beta_{k+1} E_{\mathcal{I}_m^r} E_{\mathcal{I}_m^r}^T CA\xi^*$.

Owing to $E_{\mathcal{J}_m^r}^T CA\xi^* = \mathbf{0}$, we have $\Delta z(k) = \mathbf{0}$ and $\|\Delta x(k)\| = \beta_k \|A\xi^*\| \to \infty$ as $k \to \infty$ due to $A\xi^* \neq \mathbf{0}$. It follows from Definition 1 that the system (1) with the estimator (6) is insecure.

Algorithm 2 Algorithm for the scenario $\mathcal{S}_4$

Initialization: Set $a(1) = \beta_1 \xi^*$

1: **while** $k \geq 1$ **do**
2: Set attack $b(k) = \beta_k E_{\mathcal{I}_m^r} E_{\mathcal{I}_m^r}^T CA\xi^*$
3: Set attack $a(k+1) = \beta_{k+1}\xi^* - \beta_k A\xi^*$
4: $k \leftarrow k+1$

(Necessity) We prove the necessity by contradiction. Suppose that, for any partition $\mathcal{I}_m^r = \{i_1, i_2, \cdots, i_r\}$ and $\mathcal{J}_m^r = \{j_1, j_2, \cdots, j_{m-r}\}$, one has $\mathrm{Rank}(E_{\mathcal{J}_m^r}^T CA) \geq \mathrm{Rank}(A)$, that is, $\mathrm{Rank}(E_{\mathcal{J}_m^r}^T CA) = \mathrm{Rank}(A)$.

Due to $\{a(k), b(k)\} \subseteq J(\mathcal{S}_4)$, $b(k)$ is a r-sparse vector, that is, $\exists v(k) \in \mathbb{R}^r$ such that $b(k) = E_{\mathcal{I}_m^r} v(k)$. From (19), $\Delta z(k+1) = -CA(\Delta x(k) + a(k+1)) + b(k+1)$, which implies that

$$E_{\mathcal{J}_m^r}^T \Delta z(k+1) = -E_{\mathcal{J}_m^r}^T CA(\Delta x(k) + a(k+1)) \tag{26}$$

because $E_{\mathcal{J}_m^r}^T E_{\mathcal{I}_m^r} = \mathbb{O}$ in (24). Noting that $\mathrm{Rank}(E_{\mathcal{J}_m^r}^T CA) = \mathrm{Rank}(A)$, according to Lemma 3, there exists a matrix T such that $TE_{\mathcal{J}_m^r}^T CA\xi = A\xi$ for $\forall \xi \in \mathbb{R}^n$.

According to (19), one has $\Delta x(k+1) = A[\Delta x(k) + a(k+1)] - K\Delta z(k+1) = TE_{\mathcal{J}_m^r}^T CA[\Delta x(k) + a(k+1)] - K\Delta z(k+1)$. Then, it follows from (26) that $\Delta x(k+1) = -(TE_{\mathcal{J}_m^r}^T + K)\Delta z(k+1)$ and $\|\Delta x(k+1)\| \leq \left[\|T\|\|E_{\mathcal{J}_m^r}^T\| + \|K\|\right] \|\Delta z(k+1)\|$. This implies that $\|\Delta x(k+1)\|$ is bounded if $\|\Delta z(k+1)\|$ is bounded, and the condition in the definition of insecurity is violated. The proof is now complete. ∎

Note that the proof of Theorem 3 is applicable to Theorem 2 by choosing $\mathcal{I}_m^r = \mathcal{I}_{acc}$ and $\mathcal{J}_m^r = \mathcal{I}_{pro}$ with $\mathcal{I}_{acc}$ and $\mathcal{I}_{pro}$ given in Theorem 2.

Remark 8. *To cope with the difficulties caused by the sparsity of the vector $b(k)$, a matrix splitting technique is utilized in Theorems 2-3 by constructing appropriate matrices $E_{\mathcal{I}_n^r}$ and $E_{\mathcal{J}_n^r}$ in* (25)*. The attack algorithm developed in Algorithm 2 takes advantage of $E_{\mathcal{I}_n^r}$ and $E_{\mathcal{J}_n^r}$ to design appropriate sparse attack vectors.*

Remark 9. *It can be observed from Proposition 1 and Theorems 1-3 that four different insecurity conditions under the scenarios $\mathcal{S}_i$ ($i = 1, 2, 3, 4$) are derived in the form of necessary and sufficient conditions. The sufficiency reveals the system insecurity (vulnerability) under attacks, and the necessity reflects the system robustness against attacks.*

V. FDI Attacks With Constraints On $a(k)$ and $b(k)$

In this section, we further consider two challenging yet practical scenarios $\mathcal{S}_5$ and $\mathcal{S}_6$, where the constraint on $a(k)$ is taken into account. For the convenience of discussion, we let

$$\begin{aligned} A_1 &= E_{\mathcal{I}_n^r}^T AE_{\mathcal{I}_n^r} \quad C_1 = CE_{\mathcal{I}_n^r} \\ A_2 &= E_{\mathcal{J}_n^r}^T AE_{\mathcal{I}_n^r} \quad C_2 = CE_{\mathcal{J}_n^r} \end{aligned} \tag{27}$$

with $E_{\mathcal{I}_n^r}$ and $E_{\mathcal{J}_n^r}$ defined in (25). Similar to Section III, we first discuss the case of single-FDI attacks.

A. Single-FDI attack in scenario $\mathcal{S}_5$

In this subsection, we focus on the scenario $\mathcal{S}_5$ where the single-FDI attack is launched with limited resource capacity. The corresponding attack strategy $J(\mathcal{S}_5)$ is given by $J(\mathcal{S}_5) = \{\{a(k), b(k)\} \mid a(k) \in \mathbb{R}^n$ is a r-sparse vector, $b(k) \equiv \mathbf{0}$, $k \in \mathbb{N}^+\}$.

Theorem 4. *The system* (1) *with estimator* (6) *is insecure in the attack scenario $\mathcal{S}_5$ if there exist two index sets $\mathcal{I}_n^r$ and $\mathcal{J}_n^r$ with properties in* (21) *satisfying*

$$\mathrm{Ker}(C_1A_1) \cap \mathrm{Ker}(A_2) \cap \overline{\mathrm{Ker}(A_1)} \neq \varnothing. \tag{28}$$

Proof: First, we assume that $\xi \in \mathbb{R}^r$, $\xi \in \mathrm{Ker}(C_1A_1) \cap \mathrm{Ker}(A_2) \cap \overline{\mathrm{Ker}(A_1)}$ which implies that $C_1A_1\xi = \mathbf{0}$, $A_2\xi = \mathbf{0}$ and $A_1\xi \neq \mathbf{0}$.

Algorithm 3 Algorithm for the scenario $\mathcal{S}_5$

Initialization: Set $a(1) = \beta_1 E_{\mathcal{I}_n^r}\xi$

1: **while** $k \geq 1$ **do**
2: Set attack $a(k+1) = \beta_{k+1} E_{\mathcal{I}_n^r}\xi - \beta_k E_{\mathcal{I}_n^r} A_1\xi$
3: $k \leftarrow k+1$

In Algorithm 3, $a(k)$ is a r-sparse vector due to ξ and $A_1\xi \in \mathbb{R}^r$. Noticing that $\Delta x(0) = \mathbf{0}$, it follows from Algorithm 3 and (17) that $\Delta z(k) = \mathbf{0}$ and $\Delta x(k) = \beta_k E_{\mathcal{I}_n^r} A_1\xi$ ($k \in \mathbb{N}^+$) with $\lim_{k\to\infty} \beta_k = \infty$. It can be concluded that $\|\Delta x(k)\| \to \infty$ and $\|\Delta z(k)\| = 0 \leq \alpha$ as $k \to \infty$ due to $A_1\xi \neq \mathbf{0}$. According to Definition 1, the system (1) with the estimator (6) is insecure. ∎

B. Joint-FDI attack in scenario $\mathcal{S}_6$

This subsection discusses the scenario $\mathcal{S}_6$ where general resource-constrained joint-FDI attacks are analyzed with $a(k)$ and $b(k)$ subjected to constraints. The corresponding attack strategy $J(\mathcal{S}_6)$ is given by $J(\mathcal{S}_6) = \{(a(k), b(k))|$ both $a(k) \in \mathbb{R}^n$ and $b(k) \in \mathbb{R}^m$ are confined to r-sparse vectors, $k \in \mathbb{N}^+\}$. Define two index sets $\mathcal{I}'^r_m = \{i'_1, i'_2, \cdots, i'_r\}$ and $\mathcal{J}'^r_m = \{j'_1, j'_2, \cdots, j'_{m-r}\}$ satisfying properties in (21) and let $\widetilde{C}_1 = E_{\mathcal{I}'^r_m}^T CE_{\mathcal{I}_n^r}$ and $\widetilde{C}_2 = E_{\mathcal{J}'^r_m}^T CE_{\mathcal{I}_n^r}$.

Theorem 5. *The system* (1) *with estimator* (6) *is insecure in the attack scenario $\mathcal{S}_6$ if there exist index sets $\mathcal{I}_n^r, \mathcal{J}_n^r$ and $\mathcal{I}'^r_m, \mathcal{J}'^r_m$ satisfying properties in* (21) *such that*

$$\mathrm{Ker}(\widetilde{C}_2A_1) \cap \mathrm{Ker}(A_2) \cap \overline{\mathrm{Ker}(A_1)} \neq \varnothing \tag{29}$$

with A_1 and A_2 defined in (27)*.*

Proof: First, assume that $\xi' \in \mathbb{R}^r$, $\xi' \in \mathrm{Ker}(\widetilde{C}_2A_1) \cap \mathrm{Ker}(A_2) \cap \overline{\mathrm{Ker}(A_1)}$, which implies that $A_1\xi' \neq \mathbf{0}$, $A_2\xi' = \mathbf{0}$ and $\widetilde{C}_2A_1\xi' = \mathbf{0}$. The construction procedure of $\{a(k), b(k)\}$ is presented in Algorithm 4 with $\lim_{k\to\infty} \beta_k = \infty$. Due to ξ', $A_1\xi'$ and $\widetilde{C}_1A_1\xi' \in \mathbb{R}^r$, $a(k)$ and $b(k)$ are r-sparse vectors and $\{a(k), b(k)\} \subseteq J(\mathcal{S}_6)$.

It follows from Algorithm 4 and $\Delta x(0) = \mathbf{0}$ that $\Delta x(k) = \beta_k E_{\mathcal{I}_n^r} A_1\xi'$ and $\Delta z(k) = \mathbf{0}$. As such, we have $\|\Delta z(k)\| \equiv 0$ but

Algorithm 4 Algorithm for the scenario $\mathcal{S}_6$

Initialization: Set $a(1) = \beta_1 E_{\mathcal{I}_n^r} \xi'$

1: **while** $k \geq 1$ **do**
2: Set attack $b(k) = \beta_k E_{\mathcal{I}'^r_m} \widetilde{C}_1 A_1 \xi'$
3: Set attack $a(k+1) = \beta_{k+1} E_{\mathcal{I}_n^r} \xi' - \beta_k E_{\mathcal{I}_n^r} A_1 \xi'$
4: $k \leftarrow k+1$

$\|\Delta x(k)\| \to \infty$ as $k \to \infty$, and the insecurity conditions (13)-(14) are satisfied. ■

The insecurity conditions proposed in Theorems 4-5 are sufficient but not necessary. The conditions in Theorem 5 are more relaxed than those in Theorem 4 due to $\mathrm{Ker}(C_1 A_1) \subseteq \mathrm{Ker}(\widetilde{C}_2 A_1)$. Such relaxed conditions reveal that the system is more vulnerable to joint-FDI attacks as compared with single-FDI attacks.

Remark 10. *The joint effect from sparsity of attack vectors $a(k)$ and $b(k)$ gives rise to the main difficulty in the analysis and design of joint-FDI attacks in the scenario $\mathcal{S}_6$, and such a difficulty has been specifically tackled by using a dedicated matrix splitting method. If there exist two pairs of matrices $\{E_{\mathcal{I}_n^r}, E_{\mathcal{J}_n^r}\}$ and $\{E_{\mathcal{I}'^r_m}, E_{\mathcal{J}'^r_m}\}$ satisfying* (29)*, then an explicit attack algorithm is established in Algorithm 4 to generate proper sparse attack vectors capable of driving the estimation bias to infinity but bypassing the anomaly detector.*

Remark 11. *It is worthwhile to mention that the scenario $\mathcal{S}_6$ is general and includes other possible constrained scenarios (that are not considered in this paper) as special cases. For example, one possible scenario is the joint-FDI attack with full access to the communication channels but limited resources to compromise a subset of the estimator registers. In this scenario, the corresponding insecurity condition can be derived by letting $\mathcal{I}'^r_m = \{1, 2, \cdots, m\}$ and $\mathcal{J}'^r_m = \varnothing$ in Theorem 5.*

Remark 12. *So far, the security issue has been extensively investigated in the literature for the state estimation problems of NCSs under various deception attacks, and most results have been reported from the defenders' perspective, that is, design certain algorithms to resist malicious attacks that are assumed to be of certain types according to the historical knowledge. This paper takes a different angle to look into the insecurity issue by quantifying how different FDI attacks compromise the estimation performance. Compared to existing literature, our main results exhibit the following distinctive features: 1) the proposed joint-FDI attack model is new, which is general to cover the case where the FDI attacks take place at both the communication channels and the estimator of the NCS; 2) the considered physical constraints are new that reflect the limited access as well as the resource capacity, and such constraints play a crucial role in degrading the system performance; 3) several new conditions (mostly necessary and sufficient) are established to characterize the insecurity of the system under attacks; and 4) a set of new algorithms are designed to construct the attack sequences where the sparsity in constrained attacks is specifically handled.*

VI. Simulation

Consider a discrete-time LTI system (1) with four communication channels and one remote estimator, where the system parameters are given by

$$A = \begin{bmatrix} 0.9944 & -0.1203 & -0.4302 & 1 \\ 0.0017 & 0.9902 & -0.0747 & 1 \\ 0 & 0.8187 & 0 & 1 \\ 0 & 0 & 0 & 0.5 \end{bmatrix},$$

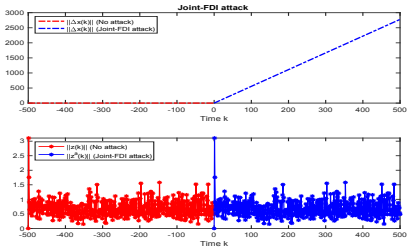

Fig. 2. The evolution of $\|\Delta x(k)\|$ and $\|z^a(k)\|$ under the joint-FDI attack in the scenario $\mathcal{S}_2$.

$C = I_4$, $Q = 0.01 I_4$ and $R = 0.1 I_3$. Then, we analyze joint-FDI attacks in different attack scenarios with different attack sequences $\{a(k), b(k)\}$.

Ideal Scenario $\mathcal{S}_2$: In the scenario $\mathcal{S}_2$, the attacker has access to all four channels. According to Algorithm 1, we let $b(k) = k[0, -0.0001, -0.0393, 0.0393]^T$ $(k \in \mathbb{N}^+)$, $a(k) = k\eta^* - (k-1)A\eta^*$ $(k \geq 2)$ with $a(1) = \eta^*$ and $\eta^* = [-0.4713, -0.1439, -0.8666, 0.0785]^T \in \mathrm{Ker}(CA)$. Under the joint-FDI attack, the estimation difference $\|\Delta x(k)\|$ goes to infinity, but the estimation residual $z^a(k)$ retains the same statistical properties with $z(k)$ (see Fig. 2).

Constrained Scenarios: Three attack scenarios $\mathcal{S}_3$, $\mathcal{S}_4$ and $\mathcal{S}_6$ for joint-FDI attacks are discussed.

Case 1: We consider the scenario $\mathcal{S}_3$ where channels 2 and 4 are under protection, and only channels 1 and 3 are accessible to the attack. In this case, the attack sequences $\{a(k), b(k)\}$ $(k \in \mathbb{N}^+)$ are generated by $a(k) = k\eta_1^* - (k-1)A\eta_1^*$ $(k \geq 2)$ with $a(1) = \eta_1^*$ and $\eta_1^* = [-0.0311, -0.0751, -0.9967, 0]^T$, and $b(k) = k[0.4069, 0, -0.0615, 0]^T$.

Case 2: We consider the scenario $\mathcal{S}_4$ where the attack has limited resources for compromising only one but one (out of four) channel(s). In this case, channel 4 is attacked and the corresponding attack sequences $\{a(k), b(k)\}$ $(k \in \mathbb{N}^+)$ are generated by $a(k) = k\eta_2^* - (k-1)A\eta_2^*$ $(k \geq 2)$ with $a(1) = \eta_2^*$ and $\eta_2^* = [0.5906, 0.3018, 0.7064, -0.2471]^T$, and $b(k) = k[0, 0, 0, -0.1236]^T$

Case 3: We consider the scenario $\mathcal{S}_6$ and the attack vectors $a(k)$ and $b(k)$ are confined to 3-sparse vectors. In this case, the attack sequences $\{a(k), b(k)\}$ $(k \in \mathbb{N}^+)$ are generated by Algorithm 4 with $\xi' = [1, 1, 1]^T \in \mathbb{R}^3$ and $\beta_k = k \in \mathbb{N}^+$.

It can be observed from Fig. 3 that, in the above three cases, the generated attack sequences have indeed driven the state estimation difference $\|\Delta x(k)\|$ to infinity with different rates while successfully evading the detection by the deployed false-data detector.

VII. Conclusion

In this paper, the security issue has been thoroughly investigated in the state estimation problems for NCSs. A novel joint-FDI attack model has been proposed to depict the situation where the false data are injected into the registers of both the communication channels and the estimator. Both limited access and resource capacity constraints have been taken into consideration in the framework of FDI attacks. Necessary and sufficient conditions for the insecurity of the system under different attack scenarios have been derived for the existence of malicious attack sequences that are capable of leading to unbounded estimation errors but bypassing the anomaly detector. Subsequently, implementable attack algorithms have been proposed to generate attack sequences over insecure NCSs for each attack scenario by resorting to the matrix splitting technique. Finally, some numerical

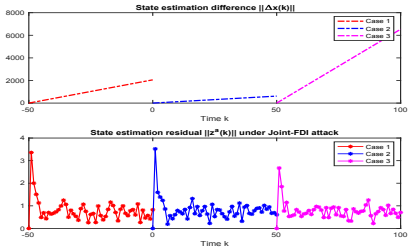

Fig. 3. The evolution of $\|\Delta x(k)\|$ and $\|z^a(k)\|$ under the joint-FDI attacks in three constrained attack scenarios. That is, case 1: the scenario $\mathcal{S}_3$ with channels 1 and 3 being attacked; case 2: the scenario $\mathcal{S}_4$ with only channel 4 being attacked; and case 3: the scenario $\mathcal{S}_6$ with 3-sparse vectors $a(k)$ and $b(k)$.

examples have been presented to verify the effectiveness of the proposed FDI attacks. Our future research topics would include 1) the security issue in the optimal control [3], [27]; 2) insecurity conditions for unsteady Kalman filters [14]; and 3) the security issue in the state estimation problem for more general systems (such as nonlinear or linear time-varying systems and wireless sensor networks) by using more sophisticated filtering algorithms [1], [2], [4], [5], [11], [14]

REFERENCES

[1] M. Basin, "Root-mean-square filtering of the state of polynomial stochastic systems with multiplicative noise," *Automation and Remote Control*, vol. 77, no. 2, pp. 242–260, 2016.

[2] M. Basin and M. Hernandez-Gonzalez, "Discrete-time H_∞ filtering for nonlinear polynomial systems," *International Journal of Systems Science*, vol. 47, no. 9, pp. 2058–2066, 2016.

[3] D. P. Bertsekas, *Dynamic programming and optimal control*. Athena Scientific Belmont, MA, 1995.

[4] R. Caballero-Aguila, A. Hermoso-Carazo and J. Linares-Perez, "Fusion estimation using measured outputs with random parameter matrices subject to random delays and packet dropouts," *Signal Processing*, vol. 127, pp. 12–23, 2016.

[5] R. Caballero-Aguila, I. Garcia-Garrido and J. Linares-Perez, "Quadratic estimation problem in discrete-time stochastic systems with random parameter matrices," *Applied Mathematics and Computation*, vol. 273, pp. 308–320, 2016.

[6] D. Ding, Q.-L. Han, Y. Xiang, X. Ge, and X. M. Zhang, "A survey on security control and attack detection for industrial cyber-physical systems," *Neurocomputing*, vol. 275, pp. 1674–1683, 2018.

[7] D. Ding, Z. Wang, D. W. C. Ho, and G. Wei, "Distributed recursive filtering for stochastic systems under uniform quantizations and deception attacks through sensor networks," *Automatica*, vol. 78, pp. 231–240, 2017.

[8] J. Farwell and R. Rohozinski, "Stuxnet and the future of cyber war," *Survival*, vol. 53, no. 1, pp. 23–40, 2011.

[9] S. Feng and P. Tesi, "Resilient control under Denial-of-Service: robust design," *Automatica*, vol. 79, pp. 42–51, 2017.

[10] J. N. Franklin, *Matrix theory*. Courier Corporation, 2012.

[11] M. J. Garcia-Ligero, A. Hermoso-Carazo and J. Linares-Perez, "Estimation from a multisensor environment for systems with multiple packet dropouts and correlated measurement noises," *Applied Mathematical Modelling*, vol. 45, pp. 802–812, 2017.

[12] X. Ge, Q.-L. Han, M. Zhong, and X. M. Zhang, "Distributed Krein space-based attack detection over sensor networks under deception attacks," *Automatica*, vol. 109, p. 108557, 2019.

[13] R. M. Gerdes, C. Winstead, and K. Heaslip, "CPS: an efficiency-motivated attack against autonomous vehicular transportation," *In Proceedings of the 29th Annual Computer Security Applications Conference*, pp. 99–108, 2013.

[14] F. Han, G. Wei, D. Ding, and Y. Song, "Local condition-based consensus filtering with stochastic nonlinearities and multiple missing measurements," *IEEE Transactions on Automatic Control*, vol. 62, no. 9, pp. 4784–4790, 2017.

[15] W. He, F. Qian, Q.-L. Han, and G. Chen, "Almost sure stability of nonlinear systems under random and impulsive sequential attacks," *IEEE Transactions on Automatic Control*, vol. 65, no. 9, pp. 3879–3886, 2020.

[16] 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 Transactions on Automatic Control*, vol. 59, no. 12, pp. 3194–3208, 2014.

[17] J. P. Hespanha, *Linear systems theory*. Princeton University Press, 2009.

[18] L. Hu, Z. Wang, Q.-L. Han, and X. Liu, "State estimation under false data injection attacks: Security analysis and system protection," *Automatica*, vol. 87, pp. 176–183, 2018.

[19] S. Kuvshinkova, "SQL slammer worm lessons learned for consideration by the electricity sector," *North American Electric Reliability Council*, vol. 1, no. 2, p. 5, 2003.

[20] F. Li and Y. Tang, "False data injection attack for cyber-physical systems with resource constraint," *IEEE Transactions on Cybernetics*, vol. 50, no. 2, pp. 729–738, 2018.

[21] Y. Li, D. Shi, and T. Chen, "False data injection attacks on networked control systems: a stackelberg game analysis," *IEEE Transactions on Automatic Control*, vol. 63, no. 10, pp. 3503–3509, 2018.

[22] G. Liang, J. Zhao, F. Luo, S. R. Weller, and Z. Y. Dong, "A review of false data injection attacks against modern power systems," *IEEE Transactions on Smart Grid*, vol. 8, no. 4, pp. 1630–1638, 2017.

[23] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Transactions on Information and System Security*, vol. 14, no. 1, pp. 1–33, 2011.

[24] L. Ma, Z. Wang, Q.-L. Han, and H. K. Lam, "Variance-constrained distributed filtering for time-varying systems with multiplicative noises and deception attacks over sensor networks," *IEEE Sensors Journal*, vol. 17, no. 7, pp. 2279–2288, 2017.

[25] Y. Mo, E. Garone, A. Casavola, and B. Sinopoli, "False data injection attacks against state estimation in wireless sensor networks," *Proceedings of the IEEE Conference on Decision and Control*, pp. 5967–5972, 2010.

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

[27] Y. Ni, Z. Guo, Y. Mo, and L. Shi, "On the performance analysis of reset attack in cyber-physical systems," *IEEE Transactions on Automatic Control*, vol. 65, no. 1, pp. 419–425, 2020.

[28] C. D. Persis and P. Tesi, "Input-to-state stabilizing control under Denial-of-Service," *IEEE Transactions on Automatic Control*, vol. 60, no. 11, pp. 2930–2944, 2015.

[29] T. Sui, Y. Mo, D. Marelli, X. M. Sun, and M. Fu, "The vulnerability of cyber-physical system under stealthy attacks," *IEEE Transactions on Automatic Control*, vol. 66, no. 2, pp. 637–650, Feb. 2021.

[30] 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.

[31] G. Wen, P. Wang, T. Huang, J. Lü, and F. Zhang, "Distributed consensus of layered multi-agent systems subject to attacks on edges," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 67, no. 9, 3152–3162, Sept. 2020.

[32] W. Xu, D. W. C. Ho, J. Zhong, and B. Chen, "Event/Self-triggered control for leader-following consensus over unreliable network with DoS attacks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 10, pp. 3137–3149, 2019.

[33] W. Xu, J. Kurths, G. Wen, and X. Yu, "Resilient event-triggered control strategies for second-order consensus," *IEEE Transactions on Automatic Control*, in press, 2022, DOI 10.1109/TAC.2021.3122382.

[34] W. Yang, Y. Zhang, G. Chen, C. Yang, and L. Shi, "Distributed filtering under false data injection attacks," *Automatica*, vol. 102, pp. 34–44, 2019.

[35] D. Ye, T. Y. Zhang, and G. Guo, "Stochastic coding detection scheme in cyber-physical systems against replay attack," *Information Sciences*, vol. 481, pp. 432-444, 2019.

[36] Y. Yuan, H. Yuan, L. Guo, H. Yang, and S. Sun, "Resilient control of networked control system under DoS attacks: A unified game approach," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 5, pp. 1786-1794, 2016.

[37] H. Zhang, P. Cheng, L. Shi, and J. Chen, "Optimal Denial-of-Service attack scheduling with energy constraint," *IEEE Transactions on Automatic Control*, vol. 60, no. 11, pp. 3023–3028, 2015.

[38] T. Y. Zhang and D. Ye, "False data injection attacks with complete stealthiness in cyberphysical systems: A self-generated approach," *Automatica*, vol. 120, p. 109117, 2020.

[39] M. Zhu and S. Martinez, "On the performance analysis of resilient networked control systems under replay attacks," *IEEE Transactions on Automatic Control*, vol. 59, no. 3, pp. 804–808, 2014.