The Vulnerability of Cyber-Physical System under Stealthy Attacks

Tianju Sui[1], Yilin Mo[2,†] Damián Marelli[3], Ximing Sun[1] and Minyue Fu[4], *Fellow IEEE*

***Abstract*—In this paper, we study the impact of stealthy attacks on the Cyber-Physical System (CPS) modeled as a stochastic linear system. An attack is characterised by a malicious injection into the system through input, output or both, and it is called stealthy (resp. strictly stealthy) if it produces bounded changes (resp. no changes) in the detection residue. Correspondingly, a CPS is called vulnerable (resp. strictly vulnerable) if it can be destabilized by a stealthy attack (resp. strictly stealthy attack). We provide necessary and sufficient conditions for the vulnerability and strictly vulnerability. For the invulnerable case, we also provide a performance bound for the difference between healthy and attacked system. Numerical examples are provided to illustrate the theoretical results.**

I. Introduction

Cyber-Physical Systems (CPSs), such as sensor networks, smart grids and transportation systems, are widely used in applications. Such a system combines a physical system with network technology to greatly improve the efficiency of the system. However, at the same time, this combination increases the vulnerability of the system. In particular, CPSs are subject to possible cyber attacks. At the physical system level, such attacks are characterised by malicious injections into the system through input, output or both.

The Stunex attack is one of the most famous CPS attack till now [1]. In June 2010, a targeted virus was injected into the Bushehr nuclear power plant through a USB flash disk. The virus replaced the measurement data from the centrifuges by a sequence of "normal" data to mislead the fault detection system to trust that the system was operating normally. Then, the virus injected input signals to accelerate the centrifuges to self destruction. This incident was reported to have caused a series of disastrous effects and destroyed over 3000 centrifuges [1]. The attacks like Stunex may penetrate the traditional information protection framework(such as FDI) of CPS.

Other examples of CPS attacks include: the Maroochy water breach [2], the blackout in brazil power grid [3], the SQL Slammer attack in Davis-Besse nuclear power plant [4], and many other industry security incidents [5]. According to the statistical data from ICS-CERT (see https://ics-cert.us-cert.gov), there were 245 CPS attacks confirmed in 2014 and the number increased to 295 in 2015.

CPS security has attracted many researchers to focus on this area [6]. The traditional efforts, such as robust statistics [7] and robust control [8], are designed to withstand certain types of failures. The popular Fault Detection and Isolation (FDI) method assumes that the failure is spontaneous [9], [10], [11]. However, CPS attacks are usually purposely designed to be stealthy and destructive, and are often done with the full or partial knowledge of the system dynamic model. Thus, it is insufficient to rely on robust control or FDI against CPS attacks. As shown in [12], an attacker can take advantage of the configuration of a power system to launch such attacks to successfully bypass the existing techniques for bad measurement detection.

For a CPS with a linear dynamic model for the physical system, many studies have been done in the detection and analysis of malicious attacks. The work of [13] studied the performance of an average consensus algorithm when individual agents in a networked system are under attack. In [14], the authors studied the detectability of attacks and pointed out that, for a noiseless system, the only undetectable space for attacks is due to the unknown initial state. In [15], an algorithm was offered to detect attacked sensors in a multi-sensors network. The work of [16] analyzed the performance of an attacked system and studied the stabilization problem using state feedback.

The above works all assumed that the physical system is noiseless, which is very restrictive. System noises would give a shelter for attacks because they may be mistaken for noises. For static systems subject to noises, [17] utilized a general evaluation standard to study the robustness of the network cluster mechanism against attacks; In [18], the authors considered the estimation problem in a smart grid and studied how does an undetectable attack change the state of the system.

For dynamic systems subject to noises, [19] studied the performance of Kalman Filter under attacks. They further studied the attack strategy and calculated the miss/false alarm

[1]Tianju Sui and Ximing Sun are with the Key Laboratory of Intelligent Control and Optimization for Industrial Equipment (Dalian University of Technology), Ministry of Education, and with the School of Control Science and Engineering, Dalian University of Technology, Dalian, China. Email: suitj@dlut.edu.cn; sunxm@dlut.edu.cn.

[2]Yilin Mo is with the Department of Automation and BNRist, Tsinghua University, Beijing, China. Email: ylmo@tsinghua.edu.cn.

[3]Damián Marelli is with the School of Automation, Guangdong University of Technology, Guangzhou, China, and with the French Argentine International Center for Information and Systems Sciences, National Scientific and Technical Research Council, Argentina. Email: Damian.Marelli@newcastle.edu.au.

[4]Minyue Fu is with the School of Electrical Engineering and Computer Science, University of Newcastle, Callaghan, NSW 2308, Australia. He also holds an Qianren Professorship at the School of Automation, Guandong University of Technology, China. Email: Minyue.Fu@newcastle.edu.au.

This work was supported by the National Natural Science Foundation of China (61803068, 61633014, 61803101 and U1701264), the National Key Research and Development Program of China (2018YFB170100, 2018AAA0101601), the Fundamental Research Funds for the Central Universities (DUT19RC(3)027) and Argentinean Agency for Scientific and Technological Promotion(PICT- 201-0985).

†Corresponding author.

rates of a χ^2 attack detector [20]. In [21], they also worked on the design of robust estimators against attacks for multi-sensors systems. Besides, [22] develop an adaptive controller that guarantees uniform ultimate boundedness of the closed-loop dynamical system in the face of adversarial sensor and actuator attacks. The works in [23], [24] extend the results to the cyber-physical systems subject to exogenous disturbances and leader-follower multiagent systems, respectively.

A lot of studies have also been done on special types of attacks. Zhang *et. al.* focused on the energy-constrained attack scheduling problem for Denial-of-Service (DoS) attacks [25]. Zhao *et. al.* studied the effect of stealthy attacks on consensus-based distributed economic dispatch [26]. Kung *et. al.* defined an ϵ-stealthy attack and analyzed its effect for scalar systems [27]. In [28], the authors worked on the multi-channel transmission schedule problem for remote state estimation under DoS attacks.

In this paper, we focus on a stochastic linear system under both sensor and actuator attacks. Firstly, we consider stealthy attacks or strictly stealthy attacks whose corresponding effect on the detection residue is either bounded or zero. It is noted that a stealthy attack is practically difficult to detect and a strictly stealthy attack is theoretically impossible to detect. We then study system's vulnerability under such attacks. A system is said to be vulnerable if it can be destabilized by a stealthy attack, or strictly vulnerable if it can be destabilized by a strictly stealthy attack. We give the necessary and sufficient conditions for both vulnerable systems and strictly vulnerable systems. To further study the performance of invulnerable system under stealthy attacks, we give a performance bound for the difference between healthy system and attacked system. These results will help to understand what kind of systems are robust to stealthy attacks and how to reduce their impact on the performance.

Focusing on a standard stochastic linear system equipped with state feedback controller and Romberg state observer, the contributions of this paper are mainly in two-folds: 1) The necessary and sufficient strict vulnerability/vulnerability conditions are given. The designers of Cyber-physical systems can check the robustness of system under stealthy attacks and understand what sensor/actuator channels are critical to the vulnerability; 2) A universal upper bound for the performance is given. The designers of Cyber-physical systems can evaluate the damage caused by attacks.

The rest of this paper is organized as follows. In Section II, we describe the models of CPS and attacks under our study. In Section III, we introduce the definitions of stealthy and strictly stealthy attacks. The notions of vulnerability and strictly vulnerability are defined according to the destabilizability of stealthy and strictly stealthy attacks. The necessary and sufficient conditions for strict vulnerability and vulnerability are given in Sections IV and V, respectively. The invulnerable system's performance bound for the difference between healthy and attacked systems is given in Section VI. In Section VII, examples are given to illustrate the theoretical results. Concluding remarks are stated in Section VIII. Some proofs are left in the Appendix.

II. Problem Formulation

A. System Model

In this paper, the Cyber-physical system is modeled as a linear discrete-time stochastic system in state-space form

$$x_{t+1} = Ax_t + Bu_t + w_t, \tag{1}$$

$$y_t = Cx_t + v_t, \tag{2}$$

where the state $x_t \in \mathbb{R}^n$, the measurement $y_t \in \mathbb{R}^m$ and the control input $u_t \in \mathbb{R}^p$. The process noise $w_t \in \mathbb{R}^n$ and the measurement noise $v_t \in \mathbb{R}^m$ obey some zero-mean stochastic distributions. Moreover, $A \in \mathbb{R}^{n\times n}$ is the system matrix, $B \in \mathbb{R}^{n\times p}$ is the actuator matrix and $C \in \mathbb{R}^{m\times n}$ is the measurement matrix. In the rest of paper, it is assumed that (A, C) is observable and (A, B) is controllable.

Furthermore, the control input u_t is assumed to be generated by a steady-state controller. To be specific, a steady-state controller is given by

$$u_t = L\hat{x}_t, \tag{3}$$

where $\hat{x}_t$ is generated by the estimator in (4) below and $L \in \mathbb{R}^{p\times n}$ is chosen such that $A + BL$ is stable.

We assume a linear time-invariant Luenberger estimator is being deployed, which has the following form:

$$\hat{x}_{t+1} = A\hat{x}_t + Bu_t + K[y_{t+1} - C(A\hat{x}_t + Bu_t)], \tag{4}$$

where K is chosen such that $A - KCA$ is stable.

Remark 1. Other than the constraint that both $A + BL$ and $A - KCA$ are stable, the choices of L and K are arbitrary.

We define the innovation signal z_t as

$$z_{t+1} = y_{t+1} - C(A\hat{x}_t + Bu_t), \tag{5}$$

and the estimation error e_t as

$$e_t = x_t - \hat{x}_t.$$

Combining (1) and (4), one can prove that e_t follows the following recursive equation:

$$e_{t+1} = (A - KCA)e_t + (I - KC)w_t - Kv_{t+1}. \tag{6}$$

B. Attack Model

In this paper, we assume that the adversary can inject an external control input and manipulate a subset of the sensory data. Therefore, system under the attack can be described by the following equations:

$$x'_{t+1} = Ax'_t + Bu'_t + B^a u^a_t + w_t, \tag{7}$$

$$y'_t = Cx'_t + \Gamma^a y^a_t + v_t, \tag{8}$$

where we use $(\cdot)'$ to denote the variable $\cdot$ under attack, $u^a_t \in \mathbb{R}^{p_a}$ is the actuator attack signal, $y^a_t \in \mathbb{R}^{m_a}$ is the sensor attack signal[1], $B^a \in \mathbb{R}^{n\times p_a}$ is the actuator attack matrix and $\Gamma^a \in \mathbb{R}^{m\times m_a} = \begin{bmatrix} e_{i_1} & \dots & e_{i_{m_a}} \end{bmatrix}$ is the sensor attack matrix, where e_i are the ith canonical basis vector of $\mathbb{R}^m$, and $\{i_1, \dots, i_{m_a}\}$

[1]In this paper, we do not put any constraint on u^a_t and y^a_t except that they need to satisfy stealthy or strictly stealthy requirement, which is introduced later in Section III.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2020.2987307, IEEE Transactions on Automatic Control

is the set of the compromised sensors. Moreover, the attack is assumed to start at time 1.

Without loss of generality, we assume that both B^a and Γ^a are full column rank.[2] The dimension p_a and m_a of the attack signal u_t^a, y_t^a represent the attacks' degrees of freedom.

In order to consider the worst-case scenario for the CPS, the attacker is assumed to know the full system model (1)-(2).

In the presence of the adversary, the steady-state estimator and controller are given by

$$\begin{aligned} \hat{x}'_{t+1} &= A\hat{x}'_t + Bu'_t + K[y'_{t+1} - C(A\hat{x}'_t + Bu'_t)], \quad (9)\\ u'_t &= L\hat{x}'_t. \end{aligned}$$

The innovation signal and estimation error are updated as

$$\begin{aligned} z'_{t+1} &= y'_{t+1} - C(A\hat{x}'_t + Bu'_t), \quad (10)\\ e'_t &= x'_t - \hat{x}'_t. \quad (11) \end{aligned}$$

The difference between an attacked system and the healthy system is characterized by

$$\begin{aligned} &\triangle x_t \triangleq x'_t - x_t, && \triangle \hat{x}_t \triangleq \hat{x}'_t - \hat{x}_t,\\ &\triangle u_t \triangleq u'_t - u_t, && \triangle y_t \triangleq y'_t - y_t,\\ &\triangle z_t \triangleq z'_t - z_t, && \triangle e_t \triangleq e'_t - e_t. \end{aligned} \quad (12)$$

The difference variables are of particular interest for the adversary and will be the focus in the rest of the paper. To be specific, Δz_t and Δy_t can be used to characterize the stealthiness of the attack. An intrusion detector employed by the CPS is unable (or hardly able) to distinguish a healthy system and a compromised system if $\triangle z_t$ and $\triangle y_t$ are zero (or small enough). The quantities $\triangle x_t$ and $\triangle e_t$ can be used to quantify the damage caused by the attack.

Remark 2. Since we assume that the attacks start at time 1, the biases between healthy and attacked system are all zeroes at time 0, i.e., $\triangle e_0 = 0, \triangle x_0 = 0$ and $\triangle z_0 = 0$.

III. Classifications for systems and attacks

In this section, we shall classify the attacks depending on the stealthiness of the attack. Since the input of any detector is the sensory data $\{y'_t : t \in \mathbb{N}\}$, and there is a one-to-one mapping between the residual error sequence $\{z'_t : t \in \mathbb{N}\}$ and the sensory data, we analyze the difference between the attacked system's z'_t and the healthy system's z_t, i.e., $\triangle z_t$, to determine if an attack can be detected or not. An attack is impossible to be detected if

$$\|\triangle z_t\| = 0, \forall t \in \mathbb{N}. \quad (13)$$

In practice, an attack is hardly detectable if $\triangle z_t$ is small enough, i.e., there exist $\delta > 0$ such that

$$\|\triangle z_t\| \leq \delta, \forall t \in \mathbb{N}. \quad (14)$$

Remark 3. As proved by Theorem 1 in [20], for a linear Gaussian system monitored by a χ^2 detector, the alarm rate converges to the false alarm rate as $\delta \to 0$. Moreover, Bai et al. [29] have proven a similar result for other forms of detectors.

Based on the above, the classification of attacks is given below.

Definition 1. An attack sequence is said to be *stealthy* if (14) is satisfied for some $\delta > 0$ and *strictly stealthy* if (13) holds.

By substracting (10) from (5) and (9) from (4), we have

$$\begin{aligned} \triangle \hat{x}_{t+1} &= (A + BL)\triangle \hat{x}_t + K\triangle z_{t+1}, \quad (15)\\ \triangle y_{t+1} &= \triangle z_{t+1} + C(A + BL)\triangle \hat{x}_t. \quad (16) \end{aligned}$$

Based on the definitions of $\triangle z_t$ and $\triangle e_t$, their update equations are given by

$$\begin{aligned} &\triangle e_{t+1}\\ &= (I - KC)A\triangle e_t + (I - KC)B^a u_t^a - K\Gamma^a y_{t+1}^a, \quad (17) \end{aligned}$$

and

$$\begin{aligned} \triangle z_{t+1} &= \triangle y_{t+1} - C(A + BL)\triangle \hat{x}_t\\ &= CA\triangle e_t + CB^a u_t^a + \Gamma^a y_{t+1}^a. \quad (18) \end{aligned}$$

Noted that both $\triangle z_t$ and $\triangle e_t$ depend only on the attack signals.

A system is resilient under the attacks if both $\limsup_{t\to\infty} \|\triangle x_t\| < \infty$ and $\limsup_{t\to\infty} \|\triangle e_t\| < \infty$. The following lemma shows that we only need to check one of the conditions instead of both.

Lemma 1. *For a stealthy or strictly stealthy attack on the system* (1)-(2)*, a necessary and sufficient condition for* $\limsup_{t\to\infty} \|\triangle x_t\| < \infty$ *is*

$$\limsup_{t\to\infty} \|\triangle e_t\| < \infty. \quad (19)$$

Proof: Based on the notations in (12), we have

$$\triangle x_t = (\hat{x}'_t + e'_t) - (\hat{x}_t + e_t) = \triangle \hat{x}_t + \triangle e_t. \quad (20)$$

Recall the equation (15), we have

$$\triangle \hat{x}_{t+1} = (A + BL)\triangle \hat{x}_t + K\triangle z_{t+1}.$$

Since $A + BL$ is stable and Δz_k is bounded due to the stealthy (or strictly stealthy) requirement in (13) and (14), the variable $\triangle \hat{x}_t$ is bounded for any $t \in \mathbb{N}$. Therefore, the condition $\limsup_{t\to\infty} \|\triangle e_t\| < \infty$ is equivalent to $\limsup_{t\to\infty} \|\triangle x_t\| < \infty$. ∎

In the rest of paper, we will use the boundness of $\limsup_{t\to\infty} \|\triangle e_t\|$ to represent the resilience under attacks. Combining with the classification of attacks in Definition 1, we can classify a system (1)-(2) depending on if there exists a stealthy (or strictly stealthy) attack to introduce an unbounded estimation error Δe_t (or bias on the state Δx_t).

Definition 2. The linear system in (1)-(2) is said to be *vulnerable* (or *strictly vulnerable*) if, for any $M_1 > 0$, there exists a *stealthy* (or *strictly stealthy*) attack such that

$$\limsup_{t\to\infty} \|\triangle e_t\| > M_1. \quad (21)$$

[2] If B^a is not full column rank, then certain column of B^a can be represented by a linear combination of other columns, i.e., the effect of certain malicious actuator on the system can be duplicated by the combined effect of several other malicious actuators. Therefore, removing the redundant actuator and corresponding column in B^a will not change the attacker's capability.

And the system is *invulnerable* (or *strictly invulnerable*) if there exists $M_2 > 0$ such that

$$\limsup_{t\to\infty} \|\triangle e_t\| \leq M_2 \tag{22}$$

for any *stealthy* (or *strictly stealthy*) attacks.

Remark 4. The vulnerability and strict vulnerability of a system are important concepts for system security. That a system is strictly invulnerable means that it is always stable under any attacks that have no influence on the residue. Meanwhile, that a system is invulnerable means that it is always stable under any attacks that have bounded influence on the residue. Thus, the invulnerable system is more robust to the attacks. We will provide necessary and sufficient conditions for vulnerability and strict vulnerability in the following Sections IV and V, respectively.

IV. The Necessary and Sufficient Condition for Strict Vulnerability

This section is devoted to the characterization of strictly vulnerable systems.

Definition 3. Consider the following linear system with initial state $x_0 = 0$:

$$x_{k+1} = Ax_k + Bu_k, y_k = Cx_k + Du_k. \tag{23}$$

The above system is said to be *invertible* if $y_k = 0$ for all $k \in \mathbb{N}$ implies that $u_k = 0$ for all $k \in \mathbb{N}$[3].

Remark 5. If a system is invertible, the mapping from the input $\{u_k : k \in \mathbb{N}\}$ to the output $\{y_k : k \in \mathbb{N}\}$ is injective, which means that different input will result in different output. In particular, any non-zero input will result in an non-zero output.

In particular, one can check the invertibility of a linear system using the following rank conditions, the proof of which can be found in [32].

Proposition 1. *The linear system in* (23) *is* invertible *if and only if*

$$\textit{rank}(M_n) - \textit{rank}(M_{n-1}) = \dim(u_k), \tag{24}$$

where

$$M_i = \begin{bmatrix} D & 0 & 0 & \cdots & 0 \\ CB & D & 0 & \cdots & 0 \\ CAB & CB & D & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ CA^{i-1}B & CA^{i-2}B & CA^{i-3}B & \cdots & D \end{bmatrix}$$

and n is the dimension of state x_k.

Furthermore, a complementary lemma is given to show the invertibility equivalence of two systems.

Lemma 2. *The system in* (23) *is not invertible if and only if*

$$x'_{k+1} = (A+KC)x'_k + (B+KD)u'_k, y'_k = Cx'_k + Du'_k. \tag{25}$$

is not invertible for any gain matrix K.

[3]The invertible system defined here is called *left invertible* in [30], [31].

Furthermore, if $\{u_k\}$ is a non-zero sequence of input for system (23) *such that $y_k = 0$ for all k, then $u'_k = u_k$ is a sequence of non-zero input for system* (25) *such that $y'_k = 0$ for all k.*

Proof: See Appendix A. ∎

Before giving the necessary and sufficient condition for strict vulnerability, we need one additional lemma:

Lemma 3. *Suppose the system* (23) *is non-invertible, and* $\begin{bmatrix} B \\ D \end{bmatrix}$ *has full column rank, then there exists a non-zero input sequence $\{u_k\}$, such that the following holds:*

$$\limsup_k \|x_k\| \to \infty, \textit{ and } y_k = 0, \forall k. \tag{26}$$

Proof: Since the system is non-invertible, there exists a non-zero sequence of input $\{u_k\}$, such that the corresponding $y_k = 0$ for all k. Without loss of generality, we shall assume that $u_0 \neq 0$, otherwise we can always trim the leading zero inputs in the sequence $\{u_k\}$. Now notice that

$$\begin{bmatrix} x_1 \\ y_0 \end{bmatrix} = \begin{bmatrix} A \\ C \end{bmatrix} x_0 + \begin{bmatrix} B \\ D \end{bmatrix} u_0 = \begin{bmatrix} B \\ D \end{bmatrix} u_0. \tag{27}$$

The fact that $\begin{bmatrix} B \\ D \end{bmatrix}$ has full column rank and $u_0 \neq 0$ implies that one of x_1 and y_0 is non-zero. Since y_k is constantly 0, we conclude that $x_1 \neq 0$. Without loss of generality, by proper scaling of u_k, we can assume that $\|x_1\| = 1$.

Now if $\limsup_k \|x_k\| \to \infty$, then we finish the proof. Otherwise, suppose that $\sup_k \|x_k\| \leq M$. We can recreate an input sequence [4] u'_k , such that

$$u'_k = \sum_{i=0}^{k} (2M+1)^{k-i} u_i. \tag{28}$$

Based on the property of linear systems, the corresponding state x'_k and measurement y'_k satisfy that

$$\begin{aligned} x'_k &= \sum_{i=0}^{k} (2M+1)^{k-i} x_i, \\ y'_k &= \sum_{i=0}^{k} (2M+1)^{k-i} y_i = 0. \end{aligned}$$

Thus, we have

$$\begin{aligned} \|x'_k\| &\geq (2M+1)^{k-1}\|x_1\| - \sum_{i=2}^{k} (2M+1)^{k-i}\|x_i\| \\ &\geq (2M+1)^{k-1} - \sum_{i=2}^{k} (2M+1)^{k-i} M \\ &= \frac{(2M+1)^{k-1}+1}{2}. \end{aligned}$$

It is obvious that $\limsup_{k\to\infty} \|x'_k\| = \infty$. ∎

We can now provide the necessary and sufficient condition for strict vulnerability:

[4]The design of u'_k utilizes the linearity combination property [33], which guarantees both $y'_k = 0$ for all $k \in \mathbb{N}$ and $x'_k = \sum_{i=0}^{k} (2M+1)^{k-i} x_k$. Moreover, the item $(2M+1)$ is used to ensure the divergence of x'_k.

Theorem 1. *The system* (1)-(2) *is strictly vulnerable if and only if the following system is not invertible:*

$$x_{k+1} = Ax_k + \begin{bmatrix} B^a & 0 \end{bmatrix} \zeta_k, \tag{29}$$
$$y_k = CAx_k + \begin{bmatrix} CB^a & \Gamma^a \end{bmatrix} \zeta_k, \tag{30}$$

where $\zeta_k = \begin{bmatrix} u_k^a \\ y_{k+1}^a \end{bmatrix}$ *is the input of system* (29)-(30).

Proof: **Sufficiency:** Firstly, based on Lemma 2, the fact that system (29)-(30) is not invertible implies that the following system is not invertible for any K:

$$\begin{aligned} x_{k+1} &= (A - KCA)x_k \\ &\quad + \begin{bmatrix} B^a - KCB^a & -K\Gamma^a \end{bmatrix} \zeta_k, \qquad (31) \\ y_k &= CAx_k + \begin{bmatrix} CB^a & \Gamma^a \end{bmatrix} \zeta_k. \qquad (32) \end{aligned}$$

Notice that

$$\begin{bmatrix} I & \\ -C & I \end{bmatrix} \begin{bmatrix} I & K \\ & I \end{bmatrix} \begin{bmatrix} B^a - KCB^a & -K\Gamma^a \\ CB^a & \Gamma^a \end{bmatrix} = \begin{bmatrix} B^a & \\ & \Gamma^a \end{bmatrix}$$

Therefore, $\begin{bmatrix} B^a - KCB^a & -K\Gamma^a \\ CB^a & \Gamma^a \end{bmatrix}$ has full column rank and by Lemma 3, there exists a sequence of ζ_k to make x_k unbounded and $y_k = 0$ for all k.

Note that (31)-(32) is an alternative expression of (17)-(18) for the dynamics of $\triangle e_t$ and $\triangle z_t$. Hence, there exists a strictly stealthy attack to make $\Delta z_t = 0$ and $\Delta e_t \to \infty$.

Necessity: Suppose the system (1)-(2) is strictly vulnerable, then there exists a non-zero input $\{\zeta_k : k \in \mathbb{N}\}$ such that $y_k = 0$ for all $k \in \mathbb{N}$ in (32). This means that the system (31)-(32) is non-invertible. Based on Lemma 2, the system (29)-(30) is also not invertible. ∎

We can further simplify our invertibility condition in Theorem 1 as follows.

Corollary 1. *The system* (1)-(2) *is strictly vulnerable if and only if the following system is not invertible:*

$$x'_{k+1} = Ax'_k + \begin{bmatrix} B^a & 0 \end{bmatrix} \zeta'_k,\ y'_k = Cx'_k + \begin{bmatrix} 0 & \Gamma^a \end{bmatrix} \zeta'_k. \tag{33}$$

Proof: We only need to prove that the system (29)-(30) is not invertible if and only if (33) is not invertible: Suppose that the input ζ_k for system (29)-(30) is of the form $\zeta_k = \begin{bmatrix} u_k^a \\ y_{k+1}^a \end{bmatrix}$. Then let

$$\zeta'_k = \begin{bmatrix} u_k^a \\ y_k^a \end{bmatrix},$$

it follows that[5]

$$x'_k = x_k,\ y'_k = y_{k-1}.$$

Therefore, the system (29)-(30) is non-invertible if and only if (33) is non-invertible. ∎

Remark 6. From the viewpoint of structured linear system, one can use Theorem 2 in [34] to derive the generic rank of the transfer function of the system described in (33) and thus check if the system is left invertible or not (i.e., if the transfer function from the input to the output is full row rank or not.)

[5]We assume that $y_{-1} = y_{-1}^a = 0$.

V. The Necessary and Sufficient Condition for Vulnerability

Next, we study the necessary and sufficient condition for vulnerability.

Definition 4. The set V is *invariant* if, for any $v \in V$, there exists u such that

$$Av + Bu \in V, Cv + Du = 0. \tag{34}$$

Let V_m be the maximum invariant subspace, the existence of which is proven in [35]. The *maximum reachable invariant set* is given by[6]

$$V^* = \text{span}\begin{pmatrix} B & AB & \dots & A^{n-1}B \end{pmatrix} \cap V_m. \tag{35}$$

A property of the maximum reachable invariant set is shown below.

Lemma 4. *If the system* (23) *is invertible, for each* $x \in V^*$*, then there exists a unique* u *such that*

$$Ax + Bu \in V^*,\ Cx + Du = 0. \tag{36}$$

Moreover, there exists a matrix Q *such that* $u = Qx$ *for every pairs* (x, u) *satisfying* (36)*.*

Proof: See Appendix B. ∎

Before the proof of necessary and sufficient condition for vulnerability, a notation is defined to facilitate Lemmas 5-7.

Definition 5. The system

$$x_{k+1} = (A + BQ)x_k,\ y_k = (C + DQ)x_k \tag{37}$$

has *unstable reachable zero-dynamic*, if there exist a vector v satisfying the following conditions:

1) v is an unstable eigenvector of $A + BQ$ and its corresponding eigenvalue is λ with $|\lambda| \geq 1$;

2) $(C + DQ)v = 0$;

3) v is reachable for (A, B).

Remark 7. The *unstable reachable zero-dynamic* in Definition 5 contains three parts:

1) The existence of zero-dynamic space for (37), which guarantees that the output (i.e., residue) is bounded;

2) The zero-dynamic space contains an unstable eigen-space of (37), which makes the state diverge;

3) The reachability of v, which implies that a vector satisfying 1) and 2) can be reached by certain sequence of input u_k.

Next, we provide a lemma on the zero-dynamic property to study the sufficiency conditions for vulnerability.

Lemma 5. *Consider the system* (23)*, for any* $M > 0$*, there exists a stealthy input sequence* $\{u_k : k \in \mathbb{N}\}$ *such that*

$$\begin{aligned} \limsup_{k\to\infty} \|x_k\| &> M, \\ \|y_k\| &\leq \delta, \forall k \in \mathbb{N}, \end{aligned}$$

[6]The invariant set V here is also called *output-nulling controlled invariant subspace* in [31] and [35]. Especially, the maximum reachable invariant set V^* is called *the maximum output-nulling controlled invariant subspace*.

if there exists a matrix Q such that the system (37) *has* unstable reachable zero-dynamic.

Proof: See Appendix C. ∎

The following two lemmas on the conditions for zero-dynamic are needed for studying the necessity conditions for vulnerability.

Lemma 6. *Suppose the system* (23) *is non-invertible, then there exists a matrix Q such that the system* (37) *has* unstable reachable zero-dynamic.

Proof: See Appendix D. ∎

Lemma 7. *Consider an invertible system* (23) *with* $\ker(B) \cap \ker(D) = \emptyset$. *Suppose that, for any $M > 0$, there exists a stealthy input sequence $\{u_k : k \in \mathbb{N}\}$ such that*

$$\begin{aligned} \limsup_{k\to\infty} \|x_k\| &> M, \\ \|y_k\| &\leq \delta, \forall k \in \mathbb{N}, \end{aligned}$$

then there exists a matrix Q such that the system (37) *has* unstable reachable zero-dynamic.

Proof: See Appendix E. ∎

Based on the results in Lemmas 5-7, the main result on vulnerability is given below.

Theorem 2. *The system* (1)-(2) *is vulnerable if and only if there exists a vector v and matrix Q satisfying the following conditions:*

1) v is an unstable eigenvector of $A + B^aQ$ and its corresponding eigenvalue is λ with $|\lambda| \geq 1$;

2) $Cv \in$ span(Γ^a);

3) v is reachable for $(A-KCA, \begin{bmatrix} B^a - KCB^a & -K\Gamma^a \end{bmatrix})$.

Proof: **Sufficiency:**. We need to show that conditions 1)-3) imply vulnerability.

Since $Cv \in$ span(Γ^a), there exist a vector y^* such that $Cv = \Gamma^a y^*$ and a matrix W such that $\lambda y^* = -Wv$. Taking a gain matrix $\begin{bmatrix} Q \\ W \end{bmatrix}$, then

$$\begin{aligned} [CA + \begin{bmatrix} CB^a & \Gamma^a \end{bmatrix} \begin{bmatrix} Q \\ W \end{bmatrix}]v &= C(A + B^aQ)v + \Gamma^a Wv \\ &= \lambda Cv - \lambda\Gamma^a y^* \\ &= 0 \end{aligned} \tag{38}$$

and

$$\begin{aligned} & [(A - KCA) + \begin{bmatrix} B^a - KCB^a & -K\Gamma^a \end{bmatrix} \begin{bmatrix} Q \\ W \end{bmatrix}]v \\ = \;& (A + B^aQ)v - K[C(A + B^aQ)v + \Gamma^a Wv] \\ = \;& \lambda v. \end{aligned} \tag{39}$$

Till now, we know that the unstable eigenvector v of $[(A - KCA) + \begin{bmatrix} B^a - KCB^a & -K\Gamma^a \end{bmatrix} \begin{bmatrix} Q \\ W \end{bmatrix}]$ satisfies the condition 1)-2) in Definition 5 for system (31)-(32) (the same as system (17)-(18) for $\triangle e_t$ and $\triangle z_t$).

Combining with that v is reachable for $(A - KCA, \begin{bmatrix} B^a - KCB^a & -K\Gamma^a \end{bmatrix})$, the condition 3) in Definition 5 is satisfied. Thus, based on the Lemma 5, the system (31)-(32) can be destabilized by a stealthy input, i.e., the system (1)-(2) is vulnerable.

Necessity: We need to show that vulnerability implies conditions 1)-3).

Note that B^a and Γ^a are both full column rank. Recall the proof of Theorem 1, we have already proved that $\begin{bmatrix} B^a - KCB^a & -K\Gamma^a \\ CB^a & \Gamma^a \end{bmatrix}$ has full column rank, i.e.,

$$\ker \begin{bmatrix} B^a - KCB^a & -K\Gamma^a \end{bmatrix} \cap \ker \begin{bmatrix} CB^a & \Gamma^a \end{bmatrix} = \emptyset$$

for any matrix K.

We will separate the rest of the proof into two cases:

1) Suppose the system (31)-(32) is non-invertible:

Following the Lemma 6, there exists a vector v and gain matrix Ψ such that

(a) v is an unstable eigenvector of the following matrix

$$A - KCA + \begin{bmatrix} B^a - KCB^a & -K\Gamma^a \end{bmatrix} \Psi. \tag{40}$$

(b) v satisfies that

$$[CA + \begin{bmatrix} CB^a & \Gamma^a \end{bmatrix} \Psi]v = 0. \tag{41}$$

(c) v is reachable for $(A-KCA, \begin{bmatrix} B^a - KCB^a & -K\Gamma^a \end{bmatrix})$.

2) Suppose that the system in (31) and (32) is invertible:

Based on the result in Lemma 7, (a)-(c) also holds.

Let $\Psi = \begin{bmatrix} Q \\ W \end{bmatrix}$, from (41), we have

$$[CA + CB^aQ + \Gamma^a W]v = 0. \tag{42}$$

Combining (42) with the fact that v is an unstable eigenvector of (40), we have

$$\begin{aligned} & [A - KCA + B^aQ - KCB^aQ - K\Gamma^a W]v \\ = \;& (A + B^aQ)v - K(CA + CB^aQ + \Gamma^a W)v \\ = \;& (A + B^aQ)v = \lambda v. \end{aligned} \tag{43}$$

Therefore, v is also an unstable eigenvector of $A + B^aQ$. Recall that (42) implies

$$\lambda Cv = -\Gamma^a Wv.$$

That is, $Cv \in$ span(Γ^a). Hence, conditions 1)-3) all hold. ∎

Remark 8. It is worth to note that the value of $\delta > 0$ in stealthy condition (14) is independent of the vulnerability condition. This is due to the linearity of the system. The adversary can always scale its attack to make Δz_t arbitrarily small, while making Δe_t diverge.

Remark 9. Following Definition 5 and Lemmas 5-7, the vulnerability condition in Theorem 2 can be divided into three parts:

1) The existence of a non-trivial *output-nulling invariant subspace* for system (31)-(32);

2) There exists a unstable eigenvector v of $(A + B^aQ)$ belonging to the above *output-nulling invariant subspace*;

3) This unstable eigenvector v is reachable for $(A - KCA, \begin{bmatrix} B^a - KCB^a & -K\Gamma^a \end{bmatrix})$.

The existence of *output-nulling invariant subspace* in 1) can be checked through the structural knowledge of linear system [34]. On the other hand, conditions 2)-3) are not

generic properties and cannot be evaluated using structural only information. Hence, the structural information about the system can provide a necessary condition on whether the system is vulnerable.

VI. A Performance Bound for Invulnerable System

Following the necessary and sufficient condition for vulnerability in Theorem 2, we understand that the bias $\triangle e_t$ between the healthy and an attacked systems is bounded when the condition is not satisfied. In this section, focusing on the invulnerable system, we will give a performance bound for the bias $\triangle e_t$ under stealthy attacks.

Recall that in (17) and (18), letting $\zeta_t = \begin{bmatrix} u_t^a \\ y_{t+1}^a \end{bmatrix}$, we have

$$\triangle z_{t+1} = CA\triangle e_t + \begin{bmatrix} CB^a & \Gamma^a \end{bmatrix} \zeta_t,$$

and

$$\triangle e_{t+1} = (I-KC)A\triangle e_t + \begin{bmatrix} (I-KC)B^a & -K\Gamma^a \end{bmatrix} \zeta_t.$$

By taking the z-transformation on both $\triangle z_{t+1}$ and $\triangle e_{t+1}$, it deduces

$$\begin{aligned} \triangle e(z) &= (zI-(I-KC)A)^{-1} \\ &\quad \cdot \begin{bmatrix} (I-KC)B^a & -K\Gamma^a \end{bmatrix} \zeta(z) \\ &= T(z)\zeta(z) \end{aligned} \tag{44}$$

and

$$\begin{aligned} \triangle z(z) &= [CA(zI-(I-KC)A)^{-1} \\ &\quad \cdot \begin{bmatrix} (I-KC)B^a & -K\Gamma^a \end{bmatrix} + \begin{bmatrix} CB^a & \Gamma^a \end{bmatrix}]\zeta(z) \\ &= S(z)\zeta(z), \end{aligned} \tag{45}$$

where $T(z) = (zI-(I-KC)A)^{-1}\begin{bmatrix} (I-KC)B^a & -K\Gamma^a \end{bmatrix}$ and $S(z) = CA(zI-(I-KC)A)^{-1}\begin{bmatrix} (I-KC)B^a & -K\Gamma^a \end{bmatrix} + \begin{bmatrix} CB^a & \Gamma^a \end{bmatrix}$.

Before the main result, we firstly introduce some new definitions to facilitate the analysis.

Definition 6. For a vector or matrix sequence $\{\xi_t : t \in \mathbb{N}\}$, we use ξ to denote the whole sequence. If ξ is a vector sequence, we define

$$\|\xi\|_{\infty,2} = \sup_{t\in\mathbb{N}} \|\xi_t\|_2,$$

and if it is a matrix sequence, we define

$$\|\xi\|_{1,\mathrm{sp}} = \sum_{t\in\mathbb{N}} \|\xi_t\|_{\mathrm{sp}},$$

where $\|\cdot\|_{\mathrm{sp}}$ denotes the spectral norm, i.e.,

$$\|\xi_t\|_{\mathrm{sp}} = \sup_{\|x\|_2=1} \|\xi_t x\|.$$

Then, a lemma is required by the proof of main result in this section.

Lemma 8. *Suppose that the system in* (1)-(2) *is invulnerable and the attack* $\{y_t^a, u_t^a : t \in \mathbb{N}\}$ *is stealthy, then we have*

$$\ker(T(z)) \supseteq \ker(S(z)), \text{ for all } |z| = 1 \tag{46}$$

and

$$\|R\|_{1,\mathrm{sp}} < \infty, \tag{47}$$

where $\|R\|_{1,sp} = \sum_{s\in\mathbb{N}} \|R_s\|_{sp}$, $R_s = \mathcal{Z}^{-1}(R(z))$, $R(z) = T(z)S^\dagger(z)$ *and* $\mathcal{Z}^{-1}(\cdot)$ *is the inverse* z*-transformation.*

Proof: See Appendix F. ∎

Based on the properties for invulnerable system in Lemma 8, a bound for $\triangle e_k$ is given in Theorem 3.

Theorem 3. *Suppose that the system in* (1)-(2) *is invulnerable and the attack* $\{y_t^a, u_t^a : t \in \mathbb{N}\}$ *is stealthy, for any* $k \in \mathbb{N}$*, we have*

$$\|\triangle e_k\|_2 \le \|R\|_{1,sp}\delta, \tag{48}$$

where δ *is the stealthy bound for the residue defined in* (14)*.*

Proof: Based on the Lemma 8, it follows from (46) that

$$\ker(T(z)) \supseteq \ker(S(z)), \text{ for all } |z| = 1.$$

It is known that $T^\dagger(z)T(z)$ is the projection onto $\ker(T(z))$ and $S^\dagger(z)S(z)$ is the projection onto $\ker(S(z))$. For that $\ker(T(z)) \supseteq \ker(S(z))$, we can pre-multiply $T^\dagger(z)T(z)$ by $S^\dagger(z)S(z)$ without changing the result.

Then, from (44) and (45), we have

$$\begin{aligned} \triangle e(z) &= T(z)\zeta(z) = T(z)T^\dagger(z)T(z)\zeta(z) \\ &= T(z)T^\dagger(z)T(z)S^\dagger(z)S(z)\zeta(z) \\ &= T(z)S^\dagger(z)(S(z)\zeta(z)) \\ &= R(z)\triangle z(z). \end{aligned} \tag{49}$$

From (49), it follows that,

$$\triangle e = TS^\dagger \triangle z = R\triangle z,$$

or equivalently,

$$\triangle e_t = \sum_{s\in\mathbb{Z}} R_s \triangle z_{t-s}.$$

Thus, combining with the boundness of $\|R\|_{1,\mathrm{sp}}$ in Lemma 8, for all $k \in \mathbb{Z}$,

$$\begin{aligned} \|\triangle e_k\|_2 &\le \sum_{s\in\mathbb{Z}} \|R_s\|_{\mathrm{sp}} \|\triangle z_{k-s}\|_2 \le \|R\|_{1,\mathrm{sp}} \|\triangle z\|_{\infty,2} \\ &\le \|R\|_{1,\mathrm{sp}}\,\delta, \end{aligned}$$

and it completes the proof. ∎

Remark 10. From (44) and (45), the quantity $T(z)S^\dagger(z)$ can be roughly viewed as a transition function from $\triangle z(z)$ to $\triangle e(z)$.

Remark 11. Since

$$\triangle\hat{x}_{t+1} = (A+BL)\triangle\hat{x}_t + K\triangle z_{t+1},$$

combining with that $A+BL$ is stable and Δz_k is bounded due to the stealthy requirement in (14), we can easily get the bound of $\|\triangle\hat{x}_t\|_2$. Thus, the bound of $\|\triangle x_t\|_2$ follows from

$$\|\triangle x_t\|_2 \le \|\triangle\hat{x}_t\|_2 + \|\triangle e_t\|_2.$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2020.2987307, IEEE Transactions on Automatic Control

VII. Simulation

In this section, numerical examples are given to verify our proposed results in Theorem 1-3. We first consider a double integrator from [20] below:

$$\begin{aligned} x_{t+1} &= \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} x_t + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u_t + B^a u_t^a + w_t, \\ y_t &= x_t + \Gamma^a y_t^a + v_t. \end{aligned} \tag{50}$$

The stationary estimator gain is given by $K = \begin{bmatrix} 0.6 & 0 \\ -1.4 & 1.6 \end{bmatrix}$. The attack is stealthy if

$$\|\triangle z_t\| \leq 1$$

for all $t \in \mathbb{N}$, and strictly stealthy if

$$\|\triangle z_t\| = 0.$$

In the rest of section, we will show that the different choices of sensor attack matrix Γ^a and actuator attack matrix B^a will make the system vulnerable, strictly vulnerable or invulnerable.

1) The attack matrices $B^a = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \Gamma^a = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$:

Following Theorem 1, the system is strictly vulnerable. Thus, a strictly stealthy attack sequence is designed by (28) and its effects on the norms of $\triangle e_t$ and $\triangle z_t$ are shown in Figure 1.

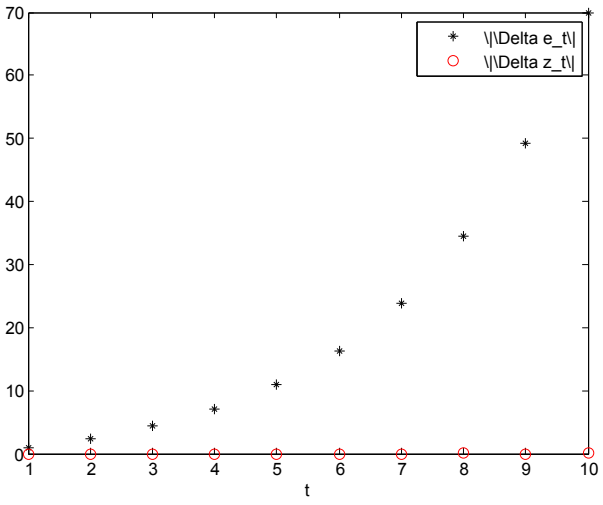

Fig. 1. The evolution of $\triangle e_t$ and $\triangle z_t$

From Figure 1, the system is destabilized by an attack while the residue bias is always zero, which confirms the strict vulnerability criterion in Theorem 1.

2) The attack matrices $B^a = 0, \Gamma^a = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$:

Following Theorems 1 and 2, the system is vulnerable but not strictly vulnerable. A stealthy attack sequence is designed by (54), and the norms of $\triangle e_t$ and $\triangle z_t$ are plotted in Figure 2.

From Figure 2, the estimation error bias between the healthy and attacked system diverges while the residual bias is kept bounded. This confirms the vulnerability criterion in Theorem 1.

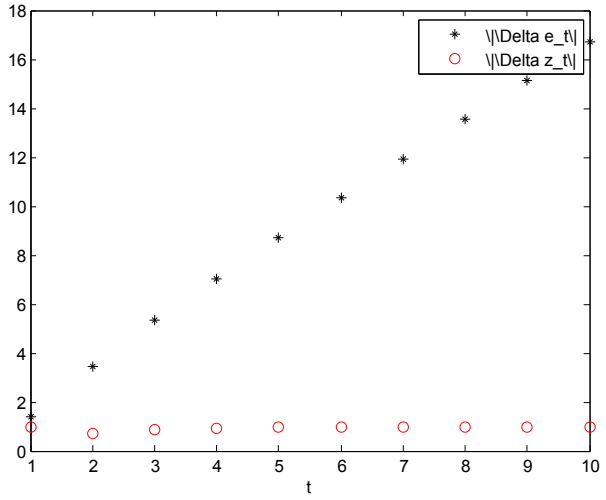

Fig. 2. The evolution of $\triangle e_t$ and $\triangle z_t$.

3) The attack matrices $B^a = 0, \Gamma^a = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$:

Following Theorem 2, the system is invulnerable for all stealthy attacks. We plot the reachable set for $\triangle e_t$ under all possible stealthy attacks in Figure 3 to show the system's robustness. Moreover, in the same figure, we also curve the universal bound for $\triangle e_t$ from Theorem 3.

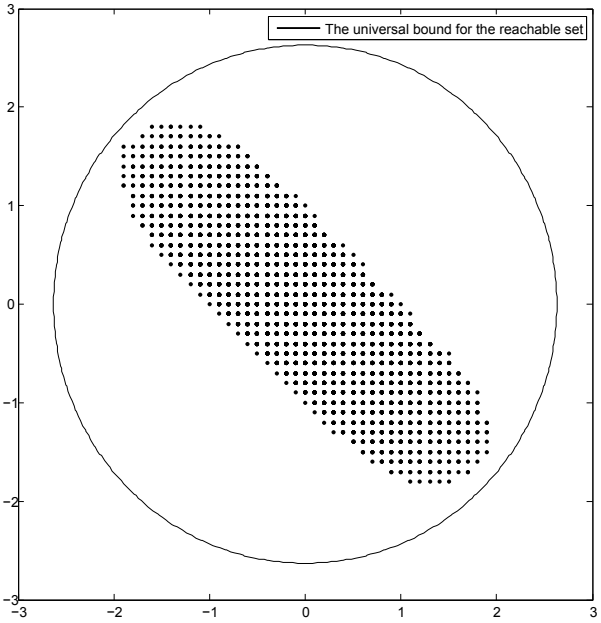

Fig. 3. The reachable set of $\triangle e_t$ and its universal bound.

The dots in Figure 3 make up the reachable set of $\triangle e_t$. From Figure 3, the estimation error bias for invulnerable system is always bounded under all stealthy attacks and its universal bound from Theorem 3 is tight and effective.

The system in (50) is simply designed to illustrate the strict/non-strict vulnerability and invulnerability properties. To show our analysis for practical system, we have introduced the well-known Tennessee Eastman Process for further simulation. Tennessee Eastman Process (TEP) is a commonly used process

proposed by Downs and Vogel in [36]. In this simulation, we adopt a simplified version of TEP from [37], as follows:

$$\begin{aligned} \dot{x} &= Ax + Bu + B^a u^a + w, \\ y &= Cx + \Gamma^a y^a + v, \end{aligned} \tag{51}$$

where A, B and C are constant matrices [7].

The TEP system is a MIMO system of order $n = 8$ with $p = 4$ inputs and $m = 10$ outputs. We discretize the system using the control system toolbox in MATLAB by selecting a sample period of 1 second. Moreover, we take $B^a = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^\top$ and $\Gamma^a = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}^\top$. Moreover, the covariance matrices Q for w and R for v are assumed to be identity matrices with proper dimensions.

Similar to that in Figure 3, we compute the 8-dimensional reachable set of the TEP model and project it onto a 2-dimensional plant and show that our universal bound from Theorem 3 is still effective.

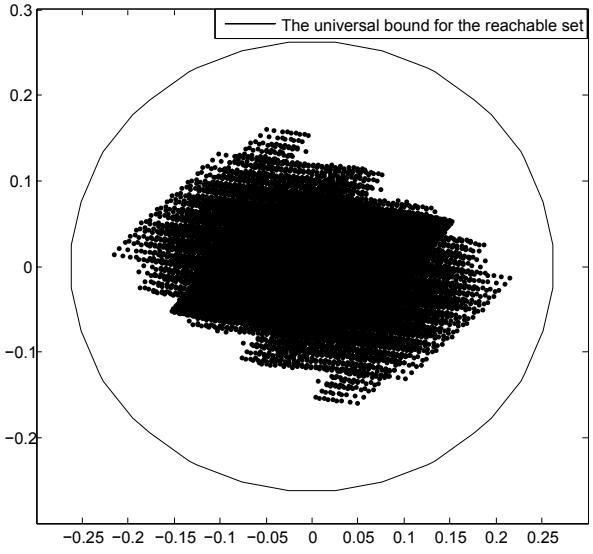

Fig. 4. The reachable set of $\triangle e_t$ and its universal bound for the TEP Model.

VIII. Conclusion

In this paper, the definitions of vulnerable and strictly vulnerable systems have been given for a stochastic linear system. A system is strictly vulnerable means that it can be destabilized by an attack that have no influence on the residue. Meanwhile, a system is vulnerable means that it can be destabilized by an attack that have bounded influence on the residue. The necessary and sufficient vulnerability and strict vulnerability conditions have been provided in this paper, which ensure the stability under stealthy and strictly stealthy attacks, respectively. Furthermore, for an invulnerable system, a performance bound for the bias between healthy and attacked system has also been given. The vulnerability condition shows what kind of system is robust to stealthy attacks and the performance bound shows how performance is affected by the stealthy attacks.

[7]For more details about this dynamic model, please refer to Appendix I in [37].

Appendix A
Proof of Lemma 2

Necessity: The system in (23) being not invertible means that there exist a nonzero input $\{u_k : k \in \mathbb{N}\}$ such that $x_0 = 0, y_k = 0$ for all $k \in \mathbb{N}$. Then, recall that in (25), we have

$$\begin{aligned} x'_{k+1} &= (A + KC)x'_k + (B + KD)u'_k \\ &= Ax'_k + Bu'_k + Ky'_k. \end{aligned} \tag{52}$$

This implies that taking $u'_k = u_k$ for all $k \in \mathbb{N}$ will make $x'_0 = 0, y'_k = 0$ for all $k \in \mathbb{N}$. Thus, the system in (25) is also not invertible.

Sufficiency: Suppose that there exist a nonzero input $\{u'_k : k \in \mathbb{N}\}$ such that $x'_0 = 0, y'_k = 0$. Recalling (52), we have

$$\begin{aligned} y'_k &= Cx'_k + Du'_k = 0, \\ x'_{k+1} &= (A + KC)x'_k + (B + KD)u'_k \\ &= Ax'_k + Bu'_k + Ky'_k = Ax'_k + Bu'_k. \end{aligned}$$

This will make $x_k = x'_k, y_k = 0$ in (25) for all k by letting $u_k = u'_k$ for all k, i.e., the system in (23) is not invertible.

Appendix B
Proof of Lemma 4

We will first prove the uniqueness of u by contradiction. Suppose there exists $u \neq u'$, such that

$$\begin{aligned} &Ax + Bu \in V^*,\, Cx + Du = 0, \\ &Ax + Bu' \in V^*,\, Cx + Du' = 0. \end{aligned}$$

By linearity of the system, we have

$$0 + B(u - u') = x_1 \in V^*, D(u - u') = 0.$$

By the property of the invariant set, there exist u_k $(k \geq 1)$ and corresponding x_k $(k \geq 2)$, such that

$$Ax_k + Bu_k = x_{k+1} \in V^*,\, Cx_k + Du_k = 0, \forall k \geq 1.$$

As a result, the non-zero control input sequence $u_0 = u - u'$, u_1, $u_2, \ldots$ results in zero output for the system, which contradicts with the assumption that the system is invertible. Thus, u satisfying (36) is unique for each $x \in V^*$.

Now we show the existence of Q. Suppose that the basis of V^* is given by $\{x_1^*, x_2^*, \ldots, x_{n^*}^*\}$, then there exists a unique set $\{u_1^*, u_2^*, \ldots, u_{n^*}^*\}$ such that

$$\begin{aligned} &Ax_i^* + Bu_i^* \in V^*, \\ &Cx_i^* + Du_i^* = 0 \end{aligned}$$

for any $i = 1, 2, \ldots, n^*$.

The matrix Q is defined as a transition matrix from $\{x_1^*, x_2^*, \ldots, x_{n^*}^*\}$ to $\{u_1^*, u_2^*, \ldots, u_{n^*}^*\}$, i.e., $u_i^* = Qx_i^*$ for any $i = 1, 2, \ldots, n^*$. Taking arbitrary $x = a_1 x_1^* + \ldots + a_{n^*} x_{n^*}^*$, we have $u = a_1 u_1^* + \ldots + a_{n^*} u_{n^*}^*$ such that

$$Ax + Bu \in V^*,\, Cx + Du = 0,$$

and $u = Qx$ always hold.

APPENDIX C
PROOF OF LEMMA 5

Recall the notations in Definition 5, since the unstable eigenvector v is reachable for (A, B), there exists $u_0, u_1, \ldots, u_{n-1}$ such that

$$x_n = v.$$

Moreover, we can manipulate the magnitude of v to satisfy that $\|y_k\| \leq \delta$ for all $k = 0, 1, \ldots, n-1$.

Then, we separate the proof into two cases:

1) Suppose $|\lambda| > 1$. A sequence of input is designed as

$$u_k = \lambda^{k-n} Q v \tag{53}$$

for any $k \geq n$.

Under the input designed above, it follows directly that $x_k = \lambda^{k-n} v$ and $y_k = 0$ for any $k \geq n$.

Hence, we have $\limsup_{k\to\infty} \|x_k\| = \infty$ and $\|y_k\| \leq \delta$ for all $k \in \mathbb{N}$.

2) Suppose $|\lambda| = 1$. A sequence of input is designed as

$$\acute{u}_{kn+j} = k\lambda^{nk+j-n} Q v + \lambda^{nk} u_j \tag{54}$$

for any $k \in \mathbb{N}$ and $j = 0, 1, \ldots, n-1$.

We derive the expression of corresponding $\acute{x}_{kn+j}$ by induction. Suppose $\acute{x}_{kn+j} = k\lambda^{nk+j-n} v + \lambda^{nk} x_j$. Then,

$$\begin{aligned}
\acute{x}_{kn+j+1} &= A\acute{x}_{kn+j} + B\acute{u}_{kn+j} \\
&= A[k\lambda^{nk+j-n} v + \lambda^{nk} x_j] \\
&\quad + B[k\lambda^{nk+j-n} Q v + \lambda^{nk} u_j] \\
&= k\lambda^{nk+j-n}(A + BQ)v + \lambda^{nk}(Ax_j + Bu_j) \\
&= k\lambda^{nk+j+1-n} v + \lambda^{nk} x_{j+1}
\end{aligned}$$

and

$$\begin{aligned}
\acute{x}_{(k+1)n} &= A\acute{x}_{kn+n-1} + B\acute{u}_{kn+n-1} \\
&= A[k\lambda^{nk-1} v + \lambda^{nk} x_{n-1}] \\
&\quad + B[k\lambda^{nk-1} Q v + \lambda^{nk} u_{n-1}] \\
&= k\lambda^{nk-1}(A + BQ)v + \lambda^{nk}(Ax_{n-1} + Bu_{n-1}) \\
&= k\lambda^{nk} v + \lambda^{nk} v \\
&= (k+1)\lambda^{(k+1)n+0-n} v + \lambda^{(k+1)n} x_0.
\end{aligned}$$

Thus, $\acute{x}_{kn+j} = k\lambda^{nk+j-n} v + \lambda^{nk} x_j$ for any $k \in \mathbb{N}$ and $j = 0, 1, \ldots, n-1$ is proved. This further implies that

$$\begin{aligned}
\acute{y}_{kn+j} &= C\acute{x}_{kn+j} + D\acute{u}_{kn+j} \\
&= C[k\lambda^{nk+j-n} v + \lambda^{nk} x_j] \\
&\quad + D[k\lambda^{nk+j-n} Q v + \lambda^{nk} u_j] \\
&= k\lambda^{nk+j-n}[C + DQ]v + \lambda^{nk}[Cx_j + Du_j] \\
&= \lambda^{nk} y_j.
\end{aligned}$$

Since $|\lambda| = 1$, we have that $\|\acute{y}_{kn+j}\| = \|y_j\| \leq \delta$ for any $k \in \mathbb{N}$ and $j = 0, 1, \ldots, n-1$. Then we conclude that $\limsup_{k\to\infty} \|\acute{x}_k\| = \infty$ and $\|\acute{y}_k\| \leq \delta$ for all $k \in \mathbb{N}$.

It is worth noting that the control input we have designed can be complex valued, as the eigenvalue and eigenvector of $A + BQ$ may be complex valued. However, by linearity, we know that if we inject the real (imaginary) part of the designed sequence u_k instead, then the state will be corresponding to the real (imaginary) part of x_k. Therefore, the divergence under a complex input means that either the real or the imaginary part of input can cause the divergence of state. Thus, we can choose the real or the imaginary part of that input as a real value input to make the system unstable.

APPENDIX D
PROOF OF LEMMA 6

Since the system in (23) is non-invertible, there exists $T \in \mathbb{N}$ and a nonzero stealthy input sequence $[u_0 = u_0^*, \ldots, u_T = u_T^*]$ with $u_0^* \neq 0$ such that $[x_1 = x_1^*, \ldots, x_T = x_T^*, x_{T+1} = a_1 x_1^* + \ldots + a_T x_T^*]$ and $y_k = 0$ for all $k \in \mathbb{N}$.

Then, with $u_0 = 0$, we have that

$$\begin{aligned}
&u_1 = 0, \ldots, u_{T-1} = 0, u_T = -a_1 u_0^* \\
&\Rightarrow x_{T+1} = -a_1 x_1^* \\
&u_1 = 0, \ldots, u_{T-1} = -a_2 u_0^*, u_T = -a_2 u_1^* \\
&\Rightarrow = x_{T+1} = -a_2 x_2^* \\
&\vdots \\
&u_1 = -a_T u_0^*, \ldots, u_{T-1} = -a_T u_{T-2}^*, u_T = -a_T u_{T-1}^* \\
&\Rightarrow x_{T+1} = -a_T x_T^*.
\end{aligned}$$

Let

$$\bar{u}_0 = u_0^*, \bar{u}_1 = u_1^* - a_T u_0^*, \ldots, \bar{u}_T = u_T^* - \sum_{i=0}^{T-1} a_{i+1} u_i^*$$

and its corresponding state sequence is denoted by

$$\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_{T+1}.$$

Based on the combination property of linear system, the input $[u_0 = \bar{u}_0, u_1 = \bar{u}_1, \ldots, u_T = \bar{u}_T]$ is nonzero and stealthy, which makes $x_{T+1} = \bar{x}_{T+1} = 0$.

For any $\lambda \geq 1$, we take

$$x = \sum_{k=0}^{T} \lambda^{-k} \bar{x}_k, \; u = \sum_{k=0}^{T} \lambda^{-k} \bar{u}_k$$

and there exists a matrix Q such that

$$u = Qx.$$

Then,

$$\begin{aligned}
(A + BQ)x &= Ax + Bu = \sum_{k=0}^{T} \lambda^{-k}(A\bar{x}_k + B\bar{u}_k) \\
&= \sum_{k=0}^{T} \lambda^{-k} \bar{x}_{k+1} = \sum_{k=1}^{T+1} \lambda^{-k+1} \bar{x}_k \\
&= \lambda \sum_{k=0}^{T} \lambda^{-k} \bar{x}_k = \lambda x
\end{aligned}$$

and

$$(C + DQ)x = Cx + Du = \sum_{k=0}^{T} \lambda^{-k}(C\bar{x}_k + D\bar{u}_k) = 0.$$

Combining with that $x = \sum_{k=0}^{T} \lambda^{-k} \bar{x}_k$ is reachable for (A, B), the conditions for the *unstable reachable zero-dynamic* of (37) in Definition 5 are satisfied and it completes the proof.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2020.2987307, IEEE Transactions on Automatic Control

APPENDIX E
PROOF OF LEMMA 7

Firstly, we will show the boundness of $\frac{\|u_k\|}{\|x_k\|+1}$.

Since the system in (23) is invertible, based on the results in Corollary 1 and Lemma 1 of [32], we have

$$u_k = \sum_{i=0}^{n-1} P_i[y_{k+i} - CA^i x_k],$$

where $P_i, i = 0, 1, \ldots, n-1$ are the gains to reconstruct the input through outputs. Thus, it follows that

$$\begin{aligned}\frac{\|u_k\|}{\|x_k\|+1} &\leq \sum_{i=0}^{n-1} \|P_i\|[\frac{\|y_{k+i}\|}{\|x_k\|+1} + \|CA^i\|\frac{\|x_k\|}{\|x_k\|+1}] \\ &\leq \sum_{i=0}^{n-1} \|P_i\|(\delta + \|CA^i\|).\end{aligned}$$

Then, there exists $U > 0$ such that $\frac{\|u_k\|}{\|x_k\|+1} \leq U$ for all $k \in \mathbb{N}$.

Denote the state x_k under stealthy input sequence $\{u_{i,t} : t \in \mathbb{N}\}$ by $x_{i,k}$. Taking arbitrary $N \in \mathbb{N}$ and $P > 0$, for the vulnerability of system (23), we could choose a cluster $\{\{u_{i,t} : t \in \mathbb{N}\} : i = 1, 2, \ldots\}$ such that[8]

$$\begin{aligned}\|x_{1,k_1}\| &> P, \\ \|x_{i+1,k_{i+1}+q}\| &> (i+1)P, \text{ for all } q = -N, \ldots, n, \\ \|x_{i+1,k_{i+1}}\| &> \max_{q=-N,\ldots,n}(\|x_{i,k_i}\|, \|x_{i+1,k_{i+1}+q}\|)\end{aligned} \quad (55)$$

And it follows directly that $\limsup_{i\to\infty} \|x_{i,k_i+q}\| = \infty$ for any $q = -N, \ldots, n$.

Based on the Bolzano-Weierstrass Theorem [38], there exists a convergent subsequence for a bounded sequence. Since

$$\|\frac{u_{i,k}}{\|x_{i,k}\|+1}\| \leq U, \|\frac{x_{i,k}}{\|x_{i,k}\|+1}\| \leq 1, \forall k \in \mathbb{N},$$

there exists a subsequence of $\{i : i \in \mathbb{N}\}$, i.e., $\{j_i : i \in \mathbb{N}\} \subseteq \{i : i \in \mathbb{N}\}$, such that

$$\begin{aligned}\lim_{i\to\infty} \frac{x_{j_i,k_{j_i}+q}}{\|x_{j_i,k_{j_i}+q}\|+1} &= \check{x}_q, q = -N, \ldots, 0, \ldots, n \\ \lim_{i\to\infty} \frac{u_{j_i,k_{j_i}+q}}{\|x_{j_i,k_{j_i}+q}\|+1} &= \check{u}_q, q = -N, \ldots, 0, \ldots, n.\end{aligned}$$

For any $q = -N, \ldots, n-1$, we have

$$\begin{aligned}A\check{x}_q + B\check{u}_q &= \lim_{i\to\infty} \frac{Ax_{j_i,k_{j_i}+q} + Bu_{j_i,k_{j_i}+q}}{\|x_{j_i,k_{j_i}+q}\|+1} \\ &= \lim_{i\to\infty} \frac{x_{j_i,k_{j_i}+q+1}}{\|x_{j_i,k_{j_i}+q}\|+1} \\ &= \underbrace{\lim_{i\to\infty} \frac{\|x_{j_i,k_{j_i}+q+1}\|+1}{\|x_{j_i,k_{j_i}+q}\|+1}}_{c_q} \check{x}_{q+1},\end{aligned} \quad (56)$$

[8]The peak sequence in (55) can be designed by

$$i_0 = 0, i_{k+1} = \min\{j : \|x_j\| > \|x_{i_k}\|\}$$

if there exist stealthy $\{u_t : t \in \mathbb{N}\}$ such that $\limsup_{k\to\infty} \|x_k\| = \infty$

and

$$\begin{aligned}C\check{x}_q + D\check{u}_q &= \lim_{i\to\infty} \frac{Cx_{j_i,k_{j_i}+q} + Du_{j_i,k_{j_i}+q}}{\|x_{j_i,k_{j_i}+q}\|+1} \\ &= \lim_{i\to\infty} \frac{y_{j_i,k_{j_i}+q}}{\|x_{j_i,k_{j_i}+q}\|+1} \\ &= 0\end{aligned} \quad (57)$$

as $\|y_{j_i,k_{j_i}+q}\| \leq \delta$ and $\limsup_{i\to\infty} \|x_{j_i,k_{j_i}+q}\| = \infty$.

Since the state x_k is n-dimensional, the vectors $\check{x}_{-N}, \ldots, \check{x}_0, \ldots, \check{x}_n$ are linearly dependent. Then, a subspace V is designed by

$$V = \text{span}[\check{x}_{-N}, \ldots, \check{x}_0, \ldots, \check{x}_d],$$

where $0 \leq d \leq n-1$ such that $\text{span}[\check{x}_{-N}, \ldots, \check{x}_0, \ldots, \check{x}_d] = \text{span}[\check{x}_{-N}, \ldots, \check{x}_0, \ldots, \check{x}_{d+1}]$, i.e., $\check{x}_{d+1} \in V$.

Then, for any $\check{x} \in V$ and let $\check{x} = b_{-N}\check{x}_{-N} + \ldots + b_d\check{x}_d$, we have

$$\begin{aligned}&A\check{x} + B[b_{-N}\check{u}_{-N} + \ldots + b_d\check{u}_d] \\ =\ & b_{-N}[A\check{x}_{-N} + B\check{u}_{-N}] + \ldots + b_d[A\check{x}_d + B\check{u}_d] \\ =\ & b_{-N}c_{-N}\check{x}_{-N+1} + \ldots + b_dc_d\check{x}_{d+1} \in V\end{aligned}$$

and

$$\begin{aligned}&C\check{x} + D[b_{-N}\check{u}_{-N} + \ldots + b_d\check{u}_d] \\ =\ & b_{-N}[C\check{x}_{-N} + D\check{u}_{-N}] + \ldots + b_d[C\check{x}_d + D\check{u}_d] \\ =\ & 0.\end{aligned}$$

Thus the subspace V is an invariant set satisfying (34). Combining with that $\check{x}_{-N}, \ldots, \check{x}_d$ are reachable for (A, B), we have $V \subseteq V^*$.

Based on the Lemma 4, since $\check{x}_{-N}, \ldots, \check{x}_0 \in V^*$, it follows that there exists matrix Q such that

$$\check{u}_q = Q\check{x}_q, \forall q = -N, \ldots, 0. \quad (58)$$

At last, we will show that $A + BQ$ is unstable on V^* by contradiction. Suppose that $A + BQ$ is stable on V^*, for that N can be arbitrary large, there exists an integer $p \leq N$ such that

$$\|(A + BQ)^p \upsilon\| < \upsilon \quad (59)$$

for all $\upsilon \in V^*$.

From (58), it follows that

$$\begin{aligned}A\check{x}_q + B\check{u}_q &= \lim_{i\to\infty} \frac{\|x_{j_i,k_{j_i}+q+1}\|+1}{\|x_{j_i,k_{j_i}+q}\|+1}\check{x}_{q+1} \\ &= (A + BQ)\check{x}_q.\end{aligned}$$

for all $q = -p, -p+1, \ldots, -1$.

The above further implies that

$$\begin{aligned}&\lim_{i\to\infty} \frac{\|x_{j_i,k_{j_i}}\|+1}{\|x_{j_i,k_{j_i}-p}\|+1}\check{x}_0 \\ =\ & \lim_{i\to\infty} \frac{\|x_{j_i,k_{j_i}}\|+1}{\|x_{j_i,k_{j_i}-1}\|+1} \cdots \frac{\|x_{j_i,k_{j_i}-p+1}\|+1}{\|x_{j_i,k_{j_i}-p}\|+1}\check{x}_0 \\ =\ & (A + BQ)^p\check{x}_{-p}.\end{aligned} \quad (60)$$

Based on the peak property defined in (55), we have that $\lim_{k\to\infty} \frac{\|x_{j_i,k_{j_i}}\|+1}{\|x_{j_i,k_{j_i}-p}\|+1} \geq 1$.

Hence, together with $\|\check{x}_{-p}\| = \|\check{x}_0\| = 1$, the equation (60) induces that

$$\begin{aligned}\|\check{x}_{-p}\| &= \|\check{x}_0\| \\ &\leq \|\lim_{i\to\infty} \frac{\|x_{j_i,k_{j_i}}\|+1}{\|x_{j_i,k_{j_i}-p}\|+1}\check{x}_0\| \\ &= \|(A+BQ)^p\check{x}_{-p}\|.\end{aligned}$$

The inequality above contradicts with the assumption in (59), thus $A+BQ$ is unstable on V^* and one of its unstable eigenvector $v \in V^*$.

Since $v \in V^* \subseteq \text{span}\begin{bmatrix}B & AB & \dots & A^{n-1}B\end{bmatrix}$, based on the Lemma 4, the conditions for the *unstable reachable zero-dynamic* of (37) in Definition 5 are satisfied and the proof is finished.

APPENDIX F
PROOF OF LEMMA 8

Since the system in (1)-(2) is invulnerable, there exists $M > 1$ such that[9]

$$\sup_{(y_t^a,u_t^a):0\leq\|\Delta z\|_{\infty,2}\leq\delta} f(\limsup_{t\to\infty}\|\Delta e_t\|_2, \limsup_{t\to\infty}\|\Delta z_t\|_2) \leq M, \tag{61}$$

where

$$f(a,b) = \begin{cases}\frac{a}{b}, & \text{if } b>0;\\ 1, & \text{if } a=0, b=0;\\ \infty, & \text{if } a>0, b=0.\end{cases}$$

Firstly, we will prove

$$\sup_{|z|=1}\sup_{0\leq\|S(z)\mu\|_2\leq\delta} f(\|T(z)\mu\|_2, \|S(z)\mu\|_2) \leq M \tag{62}$$

by contradiction. Suppose that there exists $\omega \in [-\pi,\pi)$ and μ such that

$$\frac{\left\|T\left(e^{j\omega}\right)\mu\right\|_2}{\left\|S\left(e^{j\omega}\right)\mu\right\|_2} > M. \tag{63}$$

Let the attack input

$$\zeta_t = e^{j\omega t}\mu \text{ for all } t \in \mathbb{N}.$$

It then follows from (44) and (45) that

$$\begin{aligned}\limsup_{t\to\infty}\|\Delta e_t\|_2 &= \|T\left(e^{j\omega}\right)\mu\|_2,\\ \limsup_{t\to\infty}\|\Delta z_t\|_2 &= \|S\left(e^{j\omega}\right)\mu\|_2.\end{aligned}$$

Since the magnitude of μ will not change (63), thus we let $\|\mu\|_2$ small enough to make $\left\|S\left(e^{j\omega}\right)\mu\right\|_2 \leq \|\Delta z\|_{\infty,2} \leq \delta$.

Hence, from (63),

$$\frac{\limsup_{t\to\infty}\|\Delta e_t\|_2}{\limsup_{t\to\infty}\|\Delta z_t\|_2} = \frac{\left\|T\left(e^{j\omega}\right)\mu\right\|_2}{\left\|S\left(e^{j\omega}\right)\mu\right\|_2} > M,$$

and it contradicts with (61). Since the system is by assumption invulnerable, we must have (62) and the result in (46) is thus proved.

[9]Suppose that there exists a sequence of attack such that $\|\Delta z\|_{\infty,2} = 0$ and $0 < \|\Delta e\|_{\infty,2} < \infty$, based on the linearity, there must exists another sequence of attack such that $\|\Delta z\|_{\infty,2} = 0$ and $\|\Delta e\|_{\infty,2} = \infty$. It contradicts with invulnerability.

Then, we will separate the proof for (47) into two steps:

Step 1) Let $S^*(z)$ be the conjugate transpose of $S(z)$ and $A(z) = S(z)S^*(z)$. Let $\hat{m} \in \mathbb{N}$ be the maximum number of linearly independent vectors $w_i(z)$, $i = 1,\cdots,\hat{m}$, satisfying

$$A(z)w_i(z) = 0 \text{ for all } |z| = 1.$$

Since the entries of $A(z)$ are rational functions, by solving the above, it is easy to see that the entries of $w_i(z)$ can be rational functions. Let $\tilde{m} = m - \hat{m}$ and $v_i(z)$, $i = 1,\cdots,\tilde{m}$, be a base for $\text{span}^{\perp}\{w_i(z), i = 1,\cdots,\hat{m}\}$. It is also easy to make that the entries of $v_i(z)$ are rational functions. Let $V(z) = [v_1,\cdots,v_{\tilde{m}}]$, $W(z) = [w_1,\cdots,w_{\hat{m}}]$ and $U(z) = [V,W]$. We then have

$$A(z) = U(z)\begin{bmatrix}\tilde{A}(z) & 0\\ 0 & 0\end{bmatrix}U^*(z)$$

with $\det\tilde{A}(z) \neq 0$ for at least one $|z| = 1$. Since $\tilde{A}(z) = V^*(z)A(z)V(z)$ has rational entries, $\det\tilde{A}(z)$ is a rational function. Therefore, $\det\tilde{A}(z) \neq 0$ for almost all $|z| = 1$. It then follows that

$$\begin{aligned}S(z)S^{\dagger}(z) &= S(z)S^*(z)\left(S(z)S^*(z)\right)^{\dagger}\\ &= A(z)A^{\dagger}(z)\\ &= U(z)\begin{bmatrix}I_{\tilde{m}} & 0\\ 0 & 0\end{bmatrix}U^*(z),\end{aligned}$$

where $S^{\dagger}(z)$ denotes the Moore-Penrose pseudoinverse $S(z)$ and $I_{\tilde{m}}$ denotes the identity matrix of dimension $\tilde{m}$. Hence, the inverse z-transform $SS^{\dagger}$ of $S(z)S^{\dagger}(z)$ satisfies

$$\left\|SS^{\dagger}\right\|_{1,\text{sp}} < \infty. \tag{64}$$

Step 2) Our next step is to show that $\|R\|_{1,\text{sp}}$ is finite. We do so by contradiction. Suppose that $\|R\|_{1,\text{sp}} = \infty$. For each $t \in \mathbb{Z}$, let η_t be a vector satisfying $\|\eta_t\|_2 = 1$ and

$$\|R_t\eta_t\|_2 = \|R_t\|_{\text{sp}}. \tag{65}$$

We have

$$\begin{aligned}\infty &= \sum_{t\in\mathbb{Z}}\|R_t\|_{\text{sp}} = \sum_{t\in\mathbb{Z}}\|R_t\eta_t\|_2\\ &\leq \sum_{t\in\mathbb{Z}}\|R_t\eta_t\|_1 = \sum_{d=1}^{m}\sum_{t\in\mathbb{Z}}\left|(R_t\eta_t)_d\right|,\end{aligned}$$

where $(R_t\eta_t)_d$ is the d-th element of $R_t\eta_t$.

Hence, there exists $1 \leq d \leq m$ such that

$$\sum_{t\in\mathbb{Z}}\left|(R_t\eta_t)_d\right| = \infty.$$

Then, there exists a sequence $\sigma_t \in \{-1,1\}$, $t \in \mathbb{Z}$, satisfying

$$\sum_{t\in\mathbb{Z}}\sigma_t\left(R_t\eta_t\right)_d = \infty. \tag{66}$$

Given any T, let $\xi_t = \sigma_{T-t}\eta_{T-t}$ for all $t \in \mathbb{Z}$ and $S^{\dagger} : L_2^m(\mathbb{Z}) \to L_2^{m_a+p_a}(\mathbb{Z})$ denote the Moore-Penrose pseudoinverse of S. Put

$$\zeta = S^{\dagger}\xi,$$

then

$$\Delta e = T\zeta = R\xi,$$
$$\Delta z = S\zeta = SS^{\dagger}\xi.$$

Using (66) we get

$$\begin{aligned}\lim_{T\to\infty}\|\Delta e_T\|_2 &= \lim_{T\to\infty}\left\|\sum_{t=-T}^{T} R_t\xi_{T-t}\right\|_2 \\ &= \lim_{T\to\infty}\left\|\sum_{t=-T}^{T}\sigma_t R_t\eta_t\right\|_2 = \infty. \quad (67)\end{aligned}$$

Also, using (64), for all $t\in\mathbb{Z}$,

$$\begin{aligned}\|\Delta z_t\|_2 &= \left\|\left(SS^{\dagger}\xi\right)_t\right\|_2 \le \sum_{s\in\mathbb{Z}}\left\|\left(SS^{\dagger}\right)_s\right\|_{\mathrm{sp}}\|\xi_{t-s}\|_2 \\ &\le \left\|SS^{\dagger}\right\|_{1,\mathrm{sp}}\|\xi\|_{\infty,2} < \infty. \quad (68)\end{aligned}$$

From (67) and (68), the system is vulnerable. Since by assumption the system is invulnerable, we must then have $\|R\|_{1,\mathrm{sp}} < \infty$.

REFERENCES

[1] J. Farwell and R. Rohozinski. Stuxnet and the future of cyber war. *Survival*, 53(1):23–40, 2011.

[2] J. Slay and M. Miller. Lessons learned from the maroochy water breach. *in Proceeding of Critical Infrastructure Protection*, 253:73–82, 2007.

[3] J. Conti. The day the samba stopped. *Engineering & Technology*, 5(6):46–47, 2010.

[4] S. Kuvshinkova. Sql slammer worm lessons learned for consideration by the electricity sector. *North American Electric Reliability Council*, 1(2):5, 2003.

[5] G. Richards. Hackers vs slackers. *Engineering & Technology*, 3(19):40–43, 2008.

[6] A. Cardenas, S. Amin, and S. Sastry. Research challenges for the security of control systems. In *HotSec*, 2008.

[7] P. Huber. *Robust statistics*. Springer Berlin Heidelberg, 2011.

[8] K. Zhou, J. Doyle, and K.Glover. *Robust and optimal control*. New Jersey: Prentice hall, 1996.

[9] A. Willsky. A survey of design methods for failure detection in dynamic systems. *Automatica*, 12(6):601–611, 1976.

[10] M. Massoumnia, G. Verghese, and A. Willsky. Failure detection and identification. *IEEE Transactions on Automatic Control*, 34(3):316–321, 1989.

[11] I. Hwang, S. Kim, Y. Kim, and C. Seah. A survey of fault detection, isolation, and reconfiguration methods. *IEEE Transactions on Control System Technology*, 18(3):636–653, 2010.

[12] Y. Liu, P. Ning, and M. Reiter. False data injection attacks against state estimation in electric power grids. *ACM Transactions on Information and System Security*, 14(1):1–16, 2011.

[13] F. Pasqualetti, A. Bicchi, and F. Bullo. Consensus computation in unreliable networks: A system theoretic approach. *IEEE Transactions on Automatic Control*, 57(1):90–104, 2010.

[14] F. Pasqualetti, F. Dorfler, and F. Bullo. Attack detection and identification in cyber-physical systems. *IEEE Transactions on Automatic Control*, 58(11):2715–2729, 2013.

[15] S. Sundaram, M. Pajic, C. Hadjicostis, R. Mangharam, and G. Pappas. The wireless control network: Monitoring for malicious behavior. *in Proceedings of IEEE Conference on Decision and Control*, pages 5979–5984, 2010.

[16] H. Fawzi, P. Tabuada, and S. Diggavi. Security for control systems under sensor and actuator attacks. *in Proceedings of IEEE Conference on Decision and Control*, pages 3412–3417, 2012.

[17] D. Wagner. Resilient aggregation in sensor networks. *in ACM Workshop on Security of Ad Hoc and Sensor Networks*, pages 78–87, 2004.

[18] Y. Liu, P. Ning, and M. Reiter. False data injection attacks against state estimation in electric power grids. *in Proceedings of the 16th ACM Conference on Computer and Communications Security*, 14(1):13, 2011.

[19] Y. Mo, E. Garone, A. Casavola, and B Sinopoli. False data injection attacks against state estimation in wireless sensor networks. In *49th IEEE Conference on Decision and Control*, pages 5967–5972, 2010.

[20] Y. Mo and B. Sinopoli. On the performance degradation of cyber-physical systems under stealthy integrity attacks. *IEEE Transactions on Automatic Control*, 61(9):2618–2624, 2016.

[21] Y. Mo and B. Sinopoli. Secure estimation in the presence of integrity attacks. *IEEE Transactions on Automatic Control*, 60(4):1145–1151, 2015.

[22] X. Jin, W. Haddad, and T. Yucelen. An adaptive control architecture for mitigating sensor and actuator attacks in cyber-physical systems. *IEEE Transactions on Automatic Control*, 62(11):6058–6064, 2017.

[23] X. Jin, W. Haddad, and T. Hayakawa. An adaptive control architecture for cyber-physical system security in the face of sensor and actuator attacks and exogenous stochastic disturbances. *Cyber-Physical Systems*, 4(1):39–56, 2018.

[24] X. Jin and W. Haddad. An adaptive control architecture for leader-follower multiagent systems with stochastic disturbances and sensor and actuator attacks. *International Journal of Control*, pages 1–10, 2018.

[25] H. Zhang, P. Cheng, L. Shi, and J. Chen. Optimal denial-of-service attack scheduling with energy constraint. *IEEE Transactions on Automatic Control*, 60(11):3023–3028, 2015.

[26] C. Zhao, J. He, P. Cheng, and J. Chen. Analysis of consensus-based distributed economic dispatch under stealthy attacks. *IEEE Transactions on Industrial Electronics*, 64(6):5107–5117, 2017.

[27] E. Kung, S. Dey, and L. Shi. The performance and limitations of epsilon-stealthy attacks on higher order systems. *IEEE Transactions on Automatic Control*, 62(2):941–947, 2017.

[28] K. Ding, Y. Li, D. Quevedo, S. Dey, and L. Shi. A multi-channel transmission schedule for remote state estimation under dos attacks. *Automatica*, 78:194–201, 2017.

[29] G. Bai, F. Pasqualetti, and V. Gupta. Security in stochastic control systems: Fundamental limitations and performance bounds. In *Proceedings of IEEE Conference on Decision and Control*, pages 195–200, 2015.

[30] W. Wonham. *Linear Multivariable Control: A Geometric Approach*. Springer-Verlag, New York, 1985.

[31] G. Marro, F. Morbidi, L. Ntogramatzidis, and D. Prattichizzo. Geometric control theory for linear systems: a tutorial. In *Proceedings of the 19th International Symposium on Mathematical Theory of Networks and Systems*, pages 1579–1590, 2010.

[32] M. Sain and J. Massey. Invertibility of linear time-invariant dynamical systems. *IEEE Transactions on Automatic Control*, 14(2):141–149, 1969.

[33] C. Chen. *Linear system theory and design*. Oxford University Press, Inc., 1998.

[34] J. Dion, C. Commault, and J. Woude. Generic properties and control of linear structured systems: a survey. *Automatica*, 39:1125–1144, 2003.

[35] B. Anderson. Output-nulling invariant and controllability subspace. *IFAC Proceedings Volumes*, 8(1):337–345, 1975.

[36] J. Downs and E. Vogel. A plant-wide industrial process control problem. *Computers and chemical engineering*, 17(3):245–255, 1993.

[37] L. Ricker. Model predictive control of a continuous, nonlinear, two-phase reactor. *Journal of Process Control*, 3(2):109–123, 1993.

[38] R. Bartle and D. Sherbert. *Introduction to Real Analysis(3rd ed.)*. New York: J. Wiley, 2000.

Tianju Sui received B.S. and Ph.D. degrees from Zhejiang University, Hangzhou, China. in 2012 and 2017, respectively. He is currently serving as an Associate Professor in Dalian University of Technology. His main research area includes networked estimation, distributed estimation and the security of Cyber-Physical systems.

Yilin Mo is an Associate Professor in the Department of Automation, Tsinghua University. He received his Ph.D. In Electrical and Computer Engineering from Carnegie Mellon University in 2012 and his Bachelor of Engineering degree from Department of Automation, Tsinghua University in 2007. Prior to his current position, he was a postdoctoral scholar at Carnegie Mellon University in 2013 and California Institute of Technology from 2013 to 2015. He held an assistant professor position in the School of Electrical and Electronic Engineering at Nanyang Technological University from 2015 to 2018. His research interests include secure control systems and networked control systems, with applications in sensor networks and power grids.

Damian Marelli received his Bachelors Degree in Electronics Engineering from the Universidad Nacional de Rosario, Argentina in 1995. He also received his Bachelor (Honous) degree in Mathematics and Ph.D. degree in Electrical Engineering, both from the University of Newcastle, Australia in 2003. From 2004 to 2005 he was postdoc at the Laboratoire d'Analyse Topologie et Probabilites, CNRS/Universite de Provence, France. From 2005 to 2015 he was Research Academic at the Centre for Complex Dynamic Systems and Control, the University of Newcastle, Australia. In 2007 he received a Marie Curie Postdoctoral Fellowship, hosted at the University of Vienna, and in 2011 he received a Lise Meitner Senior Fellowship, hosted at the Austrian Academy of Sciences. Since 2016, he is Professor at the School of Automation, Guangdong University of Technology, China and Independent Researcher appointment at the French-Argentinean International Center for Information and Systems Sciences, National Scientific and Technical Research Council, Argentina. His main research interests include system theory, statistical signal processing and distributed processing.

Xi-Ming Sun received the Ph.D. degree in Control Theory and Control Engineering from the Northeastern University, China, in 2006. From August 2006 to December 2008, he worked as a Research Fellow in the Faculty of Advanced Technology, University of Glamorgan, UK. He then visited the School of Electrical and Electronic Engineering, Melbourne University, Australia in 2009, and Polytechnic Institute of New York University in 2011, respectively. He is IEEE Senior Member. He serves as Associate Editor in IEEE Transactions on Cybernetics. He is currently a Professor in the School of Control Science and Engineering, Dalian University of Technology, China. He was awarded the Most Cited Article 2006-2010 from the journal of Automatica in 2011. His research interests include hybird systems, networked control systems, and nonlinear systems.

Minyue Fu (F' 03) received the B.Sc. degree in electrical engineering from the University of Science and Technology of China, Hefei, China, in 1982, and the M.S. and Ph.D. degrees in electrical engineering from the University of Wisconsin-Madison, Madison, WI, USA. in 1983 and 1987, respectively.

From 1987 to 1989, he was an Assistant Professor in the Department of Electrical and Computer Engineering, Wayne State University, USA. He joined the Department of Electrical and Computer Engineering at the University of Newcastle, Australia, in 1989. Currently, he is a Chair Professor of Electrical Engineering. He has been a Visiting Associate Professor at the University of Iowa, USA, Nanyang Technological University, Singapore and Tokyo University, Tokyo, Japan. He has held a ChangJiang Visiting Professorship at Shandong University, Jinan, China, a Extinguished Professorship at Zhejiang University and Guangdong University of Technology, China. He has been an Associate Editor for the IEEE Transactions on Automatic Control, Automatica, IEEE Transactions on Signal Processing, and the Journal of Optimization and Engineering. He is a Fellow of IEEE, Fellow of Engineers of Australia and Fellow of Chinese Association of Automation. His current research areas include networked control systems, smart electricity networks, and super-precision positioning control systems.