Revisit Input Observability

Kawano, Yu; Cao, Ming

Published in:
Proceedings of the 2018 IEEE Conference on Decision and Control (CDC)

DOI:
10.1109/CDC.2018.8618666

Revisit Input Observability: A New Approach to Attack Detection and Privacy Preservation

Yu Kawano and Ming Cao

Abstract— **Models for attack detection and privacy preservation of linear systems can be formulated in terms of their input observability, which is also called the left invertibility of their transfer function matrices. While left invertibility is a classical concept, we re-examine it from the perspectives of security and privacy. In this paper, for discrete-time linear systems, we design an input observer in order to detect attacks. We also present the input observability Gramian, which is used to characterize the systems' privacy level; it is shown that a strong connection can be made between the input observability Gramian and a standard privacy concept called differential privacy.**

I. Introduction

The Internet-of-Things (IoT) technologies have enabled fast development in smart grids [1] and health monitoring systems [2] in recent years, and the key step is the realization of remotely controlling or sensing objects through networks. While IoT is expected to dramatically change the quality of our lives, it is saddled with pressing security and privacy threats. For instance, Ukrainian power companies experienced forced power outages caused by external cyber-attacks in 2015 [3]. Researchers have recently found that when Apple implements "differential privacy" into their MacOS and iOS operating systems, the company can potentially erode the users' privacy protection [4], [5]. Motivated by these mighty events, attack detection and privacy preservation are currently being intensively studied in several fields including systems and control; for instance see [6]–[14]. In the literature, attack detection and privacy preservation problems have been investigated separately. However, we find that they can be viewed as the opposite properties of each other and thus studied in the same framework, more specifically in terms of *input observability*. For attack detection, if the attack (external inputs) can be uniquely determined, the systems can then be protected by effectively counter-acting on the attack. For privacy preservation, each individual's privacy pattern (input) should not be detected from learned results (outputs). So, input observability is a preferred property from the attack detection perspective but an undesirable property for privacy preservation.

In this paper, we focus on discrete-time linear systems, which are the common models for security and privacy analysis in systems and control [12], [13]. We refer *input observability* to the property that the initial input can be uniquely determined from the system's known initial state and measured output sequence irrespective of the choice of the input sequence. If the initial input can be uniquely determined, the whole input sequence can then be uniquely determined. Input observability is also called *invertibility with delay* in [15] and is equivalent to the *left invertibility* of the transfer function matrix [16]–[18]. Left invertibility in particular is a classical concept and can be checked by several conditions, e.g. the rank of the transfer function matrix, the PBH type test [17], [18], and Kalman's rank type conditions [15], [16]. More directly related to attack detection, input observers are provided in [17], [18], and the left invariable subspace is studied in [19], which can respectively be used to detect attacks and to identify input nodes that are vulnerable to attacks.

For attack detection, instead of the input observers, the unknown input observer (UIO) has been widely used, see, e.g. [8]–[10]. The UIO estimates the states under some unknown input which is interpreted as the attack to the system. More specifically, if there is a mismatch between an estimated state and the state computed from the state space model, then one concludes that there is an attack. Note that the UIOs do not estimate attack signals; in contrast, input observers provided by [17], [18] directly do so. However, input observers do not necessarily converge to the inputs of the original system in finite time, which can prevent them to be used for attack detection in many practical scenarios where finite-time convergence is needed. In this paper, our first goal is then to construct an input observer whose output converges to the input sequence in finite time. In fact, our input observer can be viewed as a specific left inverse system, that works even if the system's initial state is not zero, and this property does not hold for a general left inverse system.

For privacy protection, one of the most useful concepts is differential privacy [11]–[13]. It is a quantitative criterion, which has never been examined in the context of input observability. To establish a bridge between input observability and differential privacy, we extend the concept of Gramian to input observability. Like the standard observability Gramian of the initial state, the input observability Gramian can be induced from a least square estimation problem of the input, and thus a similar concept naturally appears in input estimation problems, e.g. in [14]. However, few paper has focused on the analysis of the input observability Gramian as a quantitative criterion. Based on the Gramian interpretation, in this paper, we show that the eigenvalues of the input observability Gramian can be used to evaluate the level

This work was supported in part by the European Research Council (ERC-CoG-771687) and the Dutch Organization for Scientific Research (NWO-vidi-14134).

Yu Kawano and Ming Cao are with the Faculty of Science and Engineering, University of Groningen, 9747 AG Groningen, The Netherlands. {y.kawano, m.cao}@rug.nl

of input observability especially in the context of privacy. More specifically, we clarify that differential privacy in fact evaluates the maximum eigenvalue of the input observability Gramian.

The remainder of this paper is organized as follows. Section II introduces input observability and constructs an input observer for attack detection. Section III gives the input observability Gramian in terms of which differential privacy is analyzed. In Section IV, our results are illustrated using examples from attack detection of a power network and differential privacy analysis of traffic monitoring.

II. Input Observer

Consider the following discrete-time linear system

$$\Sigma : \begin{cases} x(t+1) = Ax(t) + Bu(t), \\ y(t) = Cx(t) + Du(t), \end{cases} \tag{1}$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$ and $y \in \mathbb{R}^p$ are the state, input and output, respectively, and $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$ and $D \in \mathbb{R}^{p\times m}$. Let $U_t(i) := [u^{\mathrm{T}}(i) \ \cdots u^{\mathrm{T}}(i+t)]^{\mathrm{T}} \in \mathbb{R}^{(t+1)m}$ and $Y_t(i) := [y^{\mathrm{T}}(i) \ \cdots y^{\mathrm{T}}(i+t)]^{\mathrm{T}} \in \mathbb{R}^{(t+1)p}$ denote the input and output subsequences, respectively. For ease of notation, $x(0)$, $U_t(0)$ and $Y_t(0)$ are also written as x_0, U_t and Y_t, respectively.

For discrete-time systems, it is well known that $Y_t(i)$ can be described as a function of $x(i)$ and $U_t(i)$ namely

$$Y_t(i) = O_t x(i) + N_t U_t(i), \tag{2}$$

$$O_t := \begin{bmatrix} C^{\mathrm{T}} & (CA)^{\mathrm{T}} & \cdots & (CA^t)^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^{(t+1)p\times n},$$

$$N_t := \begin{bmatrix} D & 0 & \cdots & \cdots & 0 \\ CB & D & \ddots & & \vdots \\ CAB & CB & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & D & 0 \\ CA^{t-1}B & CA^{t-2}B & \cdots & CB & D \end{bmatrix} \in \mathbb{R}^{(t+1)p\times(t+1)m}.$$

Now, we give a formal definition of input observability studied in this paper, which is equivalent to L-delay invertibility [15], [16] and left invertibility of the transfer function matrix [17], [18] directly from their definitions.

Definition 2.1: The system Σ is said to be *input observable* for the initial state $x_0 \in \mathbb{R}^n$ if there exists a non-negative integer L such that $u(0) \in \mathbb{R}^m$ can be uniquely determined from the known initial state $x(0) = x_0$ and measured output subsequence $Y_L(0) \in \mathbb{R}^{(L+1)p}$ irrespective of the choice of the input subsequence $U_L(0) \in \mathbb{R}^{(L+1)m}$.

The reason for focusing on the initial input $u(0)$ is that one can then construct the whole input sequence from it. Actually, from $u(0)$, x_0, and system dynamics Σ, one obtains $x(1)$. Then, from $x(1)$ and $Y_L(1) \in \mathbb{R}^{(L+1)p}$, one can compute $u(1)$ and consequently $u(t)$, $t = 2, 3, \ldots$. The remaining question is an upper bound on L. Based on [15, the proof of Theorem 4], and the Cayley-Hamilton theorem [20], one can readily conclude that an upper bound is n. From the representation (2), the existence of the unique $u(0)$ can be verified as follows.

Lemma 2.2: A system Σ is input observable if and only if

$$\mathrm{rank} \begin{bmatrix} N_n^{\mathrm{T}} & e_i \end{bmatrix} = \mathrm{rank} N_n^{\mathrm{T}}, \ \forall i = 1, \ldots, m, \tag{3}$$

where $e_i \in \mathbb{R}^{(n+1)m}$ is the standard basis, i.e., its ith element is 1, and the other elements are 0.

Proof: A system Σ is input observable if and only if there exists $K \in \mathbb{R}^{m\times(n+1)p}$, not necessarily unique, such that

$$K(Y_n - O_n x_0) = K N_n U_n = u(0) \tag{4}$$

for arbitrary U_n, or equivalently, if and only if

$$K N_n = \begin{bmatrix} I_m & 0 & \cdots & 0 \end{bmatrix}.$$

A solution K exists if and only if (3) holds. ■

Our objective in this section is to design an observer that detects the attack, i.e. determining the input sequence, which we call an *input observer*. There are already several attack detectors and input observers in the literature [8]–[10], [17], [18]. The difference from them is that we aim at determining the input sequence in finite time. One can construct such an input observer if the system Σ is input observable.

Theorem 2.3: Suppose that a system Σ is input observable. Consider the following system with K satisfying (4):

$$\begin{cases} \xi(t+1) = (A - BKO_n)\xi(t) + BK\nu(t) \\ \eta(t) = -KO_n\xi(t) + K\nu(t), \end{cases} \tag{5}$$

where $\xi \in \mathbb{R}^n$, $\nu \in \mathbb{R}^{(n+1)p}$, and $\eta \in \mathbb{R}^m$ are the state, input, and output, respectively. If the initial state and input of the system (5) are chosen as $\xi(0) = x(0)$ and $\nu(t) = Y_n(t)$, then its output $\eta(t)$ is $u(t)$ for any $t = 0, 1, \ldots$.

Let $G(z)$ and $H(z)$ be the transfer function matrices of the system Σ and its input observer (5), respectively. Then, $H(z)G(z) = I_m/z^n$, i.e., $G(z)$ has the left inverse [17] $z^n H(z)$. Conversely, if the system is left invertible, there exists a transfer function $H(z)$ such that $H(z)G(z) = I_m/z^n$. For an arbitrary state space representation of $H(z)$, its output corresponding to input $Y_n(t)$ is $u(t)$ if the initial state is $x(0) = 0$. However, for non-zero initial states, this is not always true. The input observer (5) presented in this paper covers the non-zero initial state cases.

It is not clear if K satisfying (4) stabilizes $A - BKO_n$. If K is chosen such that $(A - BKO_n)^n = 0$, then the output of the input observer (5) converges to the input of the system Σ in finite time for arbitrary initial states. A matrix K satisfying $(A - BKO_n)^n = 0$ exists if the system Σ is reachable and observable. However, in general, there is no direct connection between input observability and minimality of the system Σ because the left invertibility condition is derived for minimal realization in [17]. Therefore, we remark that how to check the existence of K simultaneously satisfying (4) and achieving $(A - BKO_n)^n = 0$ is an open question.

III. Input Observability Gramians

Differential privacy [11]–[13] is known to be a quantitative criterion of privacy, and we want to establish in this section that it can be interpreted as input observability. However, for input observability, there are only qualitative (binary) criteria, such as Kalman's type rank condition [15], [16] and the PBH type condition [17], [18]. In contrast, for the standard observability of the initial state, the observability Gramian is known as a quantitative criterion. In this section, we extend the concept of Gramian to input observability, and then establish a bridge between the input observability Gramian and differential privacy.

A. *Definition of Differential Privacy*

We first provide the definition of differential privacy. The main idea of differential privacy is adding noise to the output in order to prevent the input from being determined from the output. In other words, noise is designed to make the system private, and differential privacy gives an index for designing noise.

As typically studied in differential privacy, we focus on a finite sequence of data, and thus only care about properties in finite time. That is, suppose that $u(t) = 0$, $t > M$. Consider the output with the noise $w(t) \in \mathbb{R}^p$ to be designed,

$$y^w(t) := y(t) + w(t) = Cx(t) + Du(t) + w(t). \quad (6)$$

Define $W_t := [w^{\mathrm{T}}(0) \ \cdots \ w^{\mathrm{T}}(t)]^{\mathrm{T}} \in \mathbb{R}^{p(t+1)}$ and $Y_t^w := [(y^w)^{\mathrm{T}}(0) \ \cdots \ (y^w)^{\mathrm{T}}(t)]^{\mathrm{T}} \in \mathbb{R}^{p(t+1)}$. Then, Y_t^w, $t \geq M$ can be described by

$$Y_t^w = O_t x_0 + N_{t,M} U_M + W_t, \quad (7)$$

where $N_{t,M} \in \mathbb{R}^{(t+1)p \times (M+1)m}$, $t \geq M$ ($N_{t,t} = N_t$) is

$$N_{t,M} := \begin{bmatrix} D & 0 & \cdots & 0 \\ CB & D & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ CA^{M-1}B & CA^{M-2}B & \cdots & D \\ CA^{M}B & CA^{M-1}B & \cdots & CB \\ \vdots & \vdots & & \vdots \\ CA^{t-1}B & CA^{t-2}B & \cdots & CA^{t-M-1}B \end{bmatrix}.$$

Based on the output sequence Y_t^w with noise, differential privacy can be defined. To introduce its definition, the symmetric binary relation for input sequences is still needed to be clarified. A pair of input sequences $(U_M, U'_M) \in \mathbb{R}^{(M+1)m} \times \mathbb{R}^{(M+1)m}$ is said to be $\mathrm{Adj}_2^b(U_M, U'_M)$ if

$$|U_M - U'_M|_2 \leq b \quad (8)$$

and $u(t) = u'(t)$ for any $t > M$. The differential privacy evaluates the pair of output sequences $(Y_t^w, Y_t^{w'})$ corresponding to $\mathrm{Adj}_2^b(U_M, U'_M)$ for the same initial states. If one considers the difference of the pair of the output sequences without noise, one has, with $u(t) = u'(t)$, $t > M$,

$$\begin{aligned} Y_t - Y'_t &= O_t x_0 + N_{t,t} U_t - (O_t x_0 + N_{t,t} U'_t) \\ &= N_{t,M}(U_M - U'_M). \end{aligned}$$

Therefore, to analyze a pair of outputs, one can assume $u(t) = 0$, $t > M$ and $x_0 = 0$ without loss of generality.

Now, we are ready to provide the definition of differential privacy for the system Σ.

Definition 3.1: [12], [13] The system Σ with output (6) is said to be (ε, δ)-differentially private for $\mathrm{Adj}_2^b(U_M, U'_M)$ at a finite time $t \geq M$ if there exist $\varepsilon > 0$ and $\delta \geq 0$ such that

$$\mathbb{P}(N_{t,M} U_M + W_t \in S) \leq \mathrm{e}^{\varepsilon} \mathbb{P}(N_{t,M} U'_M + W_t \in S) + \delta \quad (9)$$

for some probability distribution function $\mathbb{P} : S \to [0, 1]$ for any element S of the Borel σ-algebra on $\mathbb{R}^{p(t+1)}$, where e is Euler's number.

If ε and δ are large, the probability distribution of the output sequences $(Y_t^w, Y_t^{w'})$ corresponding to $\mathrm{Adj}_2^b(U_M, U'_M)$ are very different, which means that for a different pair of inputs, the corresponding outputs can be very different. Therefore, it is relatively easy to estimate the input sequence from the known initial state and measured output sequence, i.e., the system can be viewed as highly input observable and thus less private.

B. *Input Observability Gramians*

Relating to differential privacy, we consider a least square estimation problem of an input sequence, which naturally induces the input observability Gramian (note that the controllability Gramian is originally obtained from the minimum energy control problem [21], the dual of the least square estimation problem of the initial state).

We continue to assume that $u(t) = 0$, $t > M$ and $x_0 = 0$. For measured output sequence Y_t^w, $t \geq M$ with measurement noise, find U_M such that

$$\min_{U_M} |Y_t^w - N_{t,M} U_M|_2^2. \quad (10)$$

The least square estimation problem is well studied, and the results can be applied to (10). Define a symmetric matrix

$$\mathcal{O}_{U_M,t} := N_{t,M}^{\mathrm{T}} N_{t,M} \in \mathbb{R}^{(M+1)m \times (M+1)m}. \quad (11)$$

We call $\mathcal{O}_{U_M,t}$ in (11) the *input observability Gramian.* Note that since the input observability Gramian evaluates the input-output behavior, it does not depend on the choice of coordinates in contrast to the standard observability Gramian [22].

The least square estimation problem (10) has a unique solution if and only if $\mathcal{O}_{U_M,t}$ is non-singular, and the unique solution is

$$U_M = \mathcal{O}_{U_M,t}^{-1} N_{t,M}^{\mathrm{T}} Y_t^w. \quad (12)$$

From the structure of (12), one observes a similar property of the standard observability Gramian, which states that the eigenvectors associated with relatively large eigenvalues of $\mathcal{O}_{U_M,t}$ correspond to the set of input sequences U_M that are relatively easy to estimate (and thus less private). In fact, its maximum eigenvalue, denoted by $\lambda_{\max}(\mathcal{O}_{U_M,t})$, characterizes the differential privacy with Gaussian noise as a result of the choice of 2-norm in (10). If one considers a different noise, one needs to consider a different norm, e.g. 1-norm for Laplace noise.

Theorem 3.2: Let $W_t \sim \mathcal{N}(0, \sigma^2 I_{(t+1)p})$. Then, a system Σ with the output (6) is (ε, δ)-differentially private for $\varepsilon > 0$,

$1/2 > \delta > 0$ and $\mathrm{Adj}_2^b(U_M, U_M')$ at a finite time $t \geq M$ if σ is chosen such that

$$\sigma \geq \frac{b\lambda_{\max}^{1/2}(\mathcal{O}_{U_M,t})}{2\varepsilon}\left(\mathcal{Q}^{-1}(\delta) + \sqrt{(\mathcal{Q}^{-1}(\delta))^2 + 2\varepsilon}\right) \quad (13)$$

holds, where $\mathcal{Q}(w)$ is the so called $\mathcal{Q}$-function

$$\mathcal{Q}(w) := \frac{1}{\sqrt{2\pi}}\int_w^\infty \mathrm{e}^{-\frac{v^2}{2}}dv,$$

and $\mathcal{Q}(w) < 1/2$ for $w > 0$.

In (13), only $\lambda_{\max}(\mathcal{O}_{U_M,t})$ depends on the system Σ. Theorem 3.2 shows that if $\lambda_{\max}(\mathcal{O}_{U_M,t})$ is small, then small noise is enough to achieve (ε,δ)-differential privacy for given $\varepsilon > 0$ and $1/2 > \delta > 0$. This observation relates to the least square estimation problem (10), in the sense that if $\lambda_{\max}(\mathcal{O}_{U_M,t})$ is small, all eigenvalues of the input observability Gramian $\mathcal{O}_{U_M,t}$ are small, and solving the least square estimation problem (10) is numerically difficult, i.e. the input information is highly private. In this case, small noise is enough to protect the privacy of the input information.

To gain deeper insight, we take a further look at the eigenvalues of $\mathcal{O}_{U_M,t}$ from three aspects. First, from (11), the ith $m \times m$ block diagonal element of $\mathcal{O}_{U_M,t}$ is

$$\mathcal{O}_{u(0),t-i} = D^{\mathrm{T}}D + \sum_{k=0}^{t-i}(CA^kB)^{\mathrm{T}}CA^kB, \quad i = 1,\dots,M+1$$

where $\mathcal{O}_{u(0),-1} = D^{\mathrm{T}}D$. This is the input observability Gramian with respect to the initial input, which we call the *initial input observability Gramian*. From the relation between the eigenvalues and the trace, the sum of the eigenvalues of $\mathcal{O}_{U_M,t}$ is the sum of the eigenvalues of all $\mathcal{O}_{u(0),t-i}$, $i = 1,\dots,M+1$. Therefore, if the initial input observability Gramian has large eigenvalues, the input observability Gramian $\mathcal{O}_{U_M,t}$ has large eigenvalues either. In other words, the privacy level of the whole input sequence is characterized by that of the initial input. This is natural, since the output at each time instant contains information of the initial input, i.e. the initial input is the least private.

Next, for fixed M, $\lambda_{\max}(\mathcal{O}_{U_M,t})$ is non-decreasing with respect to t, and thus the privacy level ε in Theorem 3.2 is non-decreasing with respect to t. This corresponds to the natural observation that more data are being collected, less private a system becomes. Finally, for fixed t, the minimum eigenvalue of $\mathcal{O}_{U_M,t}$, denoted by $\lambda_{\min}(\mathcal{O}_{U_M,t})$ is not increasing with respect to M. For instance,

$$\lambda_{\min}(\mathcal{O}_{U_1,t}) \leq \lambda_{\min}(\mathcal{O}_{u(0),t}). \quad (14)$$

Recall that these two Gramians are obtained from the least square estimation problems when $u(t) = 0$ for $t = 2,3,\dots$ and $t = 1,2,\dots$, respectively. Therefore, (14) corresponds to the natural observation that $u(0)$ is more difficult to estimate if $u(1)$ is unknown compared to the case when $u(1)$ is known to be 0.

C. Input Observability Analysis

The standard controllability and observability Gramians provide not only quantitative criteria but also qualitative criteria. Here, we study the connection between the input observability Gramian and input observability.

If there is no measurement noise, i.e. $Y_t^w = Y_t$, then (12) gives exact U_M. Therefore, non-singularity of the input observability Gramian $\mathcal{O}_{U_M,t}$ is a necessary and sufficient condition for input observability when $u(t) = 0$, $t > M$. Note that this does not imply input observability for non-zero $u(t)$ in general. However, according to [16, Corollary 2], this does if $M \geq n$. Then, we have a necessary and sufficient condition for input observability.

Proposition 3.3: A system Σ is input observable if and only if $\mathcal{O}_{U_M,t}$ is non-singular for any $M \geq n$ and $t \geq M+n$.

The input observability Gramian is both a qualitative and quantitative criterion for input observability. For differential privacy, only the maximum eigenvalue is evaluated. For more detailed privacy (input observability) analysis, each eigenvalue and the associated eigen-space can be used as typically done for the standard observability Gramian. Let $v_i \in \mathbb{R}^{(2n+1)m}$, $i = 1,\dots,(2n+1)m$ be eigenvectors of $\mathcal{O}_{U_n,2n}$ associated with eigenvalues $\lambda_i \leq \lambda_{i+1}$. If there is k such that $\lambda_k \ll \lambda_{k+1}$, then $U_n \in \mathrm{span}\{v_{k+1},\dots,v_n\}$ is relatively easy to observe. Especially, if $0 < \lambda_{k+1}$, then such U_n can be uniquely determined, and the projection of $\mathrm{span}\{v_{k+1},\dots,v_n\}$ onto the $u(0)$-space gives the input observable subspace. The input observable and unobservable subspaces themselves have already been studied in [19], but quantitative analysis has not been established yet.

The quantitative analysis of subspaces can be used for designing noise to make a system more private. Let $\lambda_k \ll \lambda_{k+1}$, and consider the projection of $\mathrm{span}\{v_{k+1},\dots,v_n\}$ onto the $u(0)$-space, which we denote by $\mathcal{U} \subset \mathbb{R}^m$. Then, the output of the system is sensitive for inputs in $\mathcal{U}$. In other words, such inputs are less private. To protect less private input information, one can add noise $v \in \mathcal{U}$ to the input channels. Since the output is sensitive for inputs in $\mathcal{U}$, small input noise may be enough to protect the input information. However, differential privacy analysis is technically more involved because of the computation of probability distribution function $\mathbb{P}$; in particular, it is not always easy to find a suitable change of variables as done in the proof of Theorem 3.2.

The input observability Gramian has a strong connection with the standard observability Gramian

$$\mathcal{O}_{x,t} := \sum_{k=0}^{t-1}(CA^k)^{\mathrm{T}}(CA^k). \quad (15)$$

From these definitions (11) and (15), we have

$$\mathcal{O}_{u(0),t} = D^{\mathrm{T}}D + B^{\mathrm{T}}\mathcal{O}_{x,t}B. \quad (16)$$

If the system Σ is Schur stable, $\mathcal{O}_{x,t}$ and thus $\mathcal{O}_{u(0),t}$ are bounded for any $t \geq 0$, where t can be ∞. From the discussion about eigenvalues of $\mathcal{O}_{u(0),t}$ and $\mathcal{O}_{U_M,t}$ in the previous subsection, the input observability Gramian $\mathcal{O}_{U_M,t}$

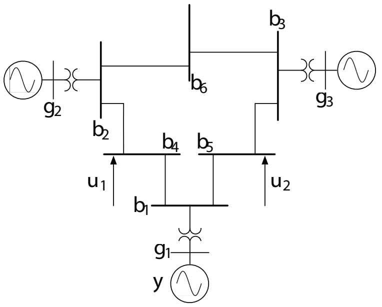

Fig. 1. WSSC power system with 3 generators and 6 buses

is bounded for any $t \geq M \geq 0$. Therefore, one can evaluate the privacy level for an infinite input sequence.

IV. Examples

A. Attack Detection for Power Networks

Consider the power network illustrated by Fig. 1, whose model can be found in [8]. We use its zero-order hold discretization with the sampling time $t = 0.01$ and consider the same B, C, and D matrices as in [8], i.e., we assume that the load buses 4 and 5 are attacked, and the monitoring unit measures the frequency and angular velocity of the first generator. In summary, we use the following model:

$$A = \begin{bmatrix} 1 & 0 & 0 & 0.01 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0.01 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0.01 \\ -0.0023 & 0.0012 & 0.0012 & 0.99 & 0 & 0 \\ 0.0044 & -0.0085 & 0.0041 & 0 & 0.98 & 0 \\ 0.0090 & 0.0087 & -0.0178 & 0 & 0 & 0.97 \end{bmatrix},$$

$$B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0.0188 & 0.0196 \\ 0.1596 & 0.0697 \\ 0.1387 & 0.3236 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix},$$

$$D = 0.$$

We verify condition (3) in Lemma 2.2. Then, $\mathrm{rank} N_6 = 6$, and condition (3) does not hold for e_1 or e_2. Therefore, attacks on load buses 4 and 5 cannot be uniquely determined. In fact, in the domain of the z-transform, $Y(z) = 0$ for $x_0 = 0$ if the input $U(z) = [U_1(z)\ U_2(z)]^{\mathrm{T}}$ satisfies

$$U_1(z) \\ = -\frac{1.962z^4 - 9.812z^3 + 21.5z^2 - 21.94z + 9.182}{1.884z^4 - 9.419z^3 + 19.75z^2 - 20.38z + 9.182} U_2(z).$$

That is, the power network is vulnerable to these attacks.

The next scenario is that the monitoring unit measures the frequencies of the first and second generators, i.e.,

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

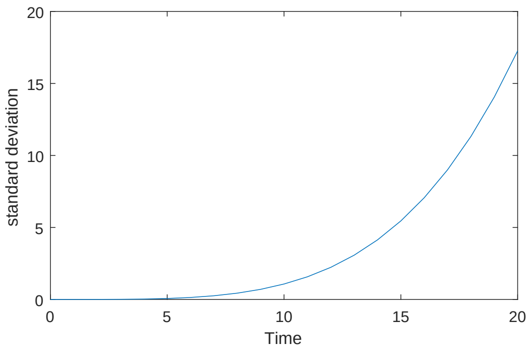

Fig. 2. Standard deviation σ of Gaussian noise to be designed in order to achieve $(0.1, 0.1)$-differential privacy

In this case, $\mathrm{rank} N_6 = 10$, and condition (3) holds for both e_1 and e_2. That is, the power network is input observable. Then, we consider to construct an input observer. For instance, K satisfying (4) is

$$K = \begin{bmatrix} 0 & 0 & 0 & 0 & -3889 & 1111 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 8889 & -1111 & 0 & \cdots & 0 \end{bmatrix}.$$

By using this K, one can construct the input observer in (5).

One notices that the gain K of the input observer is much larger than the elements of A, B, and C matrices. Therefore, one can claim that it is still difficult to detect attacks even when it is possible. To evaluate difficulty, we compute the eigenvalues of the initial input observability Gramian in (11) with $M = 0$ and $t = n = 6$. Then, its eigenvalues are 0.006×10^{-4} and 0.169×10^{-3}. As expected, they are small. This quantitative evaluation is doable thanks to our input observability formulation of an attack detection problem.

B. Differential Privacy in Traffic Monitoring

Consider a simplified traffic monitoring system studied in [12]. The purpose of the traffic monitoring service is to provide continuous estimation of the traffic flow, i.e., computing the average position of the vehicles.

Let us consider 10 vehicles whose dynamics are given by

$$x_i(t+1) = \begin{bmatrix} 1 & T_s \\ 0 & 1 \end{bmatrix} x_i(t) + \begin{bmatrix} 0 \\ T_s \end{bmatrix} u_i(t), \\ i = 1, \dots, r,$$

where $T_s = 0.01$ is a sampling period of the position measurement, $x_i = [\xi_i\ \dot{\xi}_i]^{\mathrm{T}}$ with ξ_i and $\dot{\xi}_i$ being position and velocity of vehicle i, and u_i is the acceleration input. The output is the average position of the vehicles,

$$y(t) = \frac{1}{10} \sum_{i=1}^{10} \xi_i(t).$$

The acceleration u_i of each vehicle is determined by each driver and thus contains information of the personal driving style. To protect this information, the Gaussian noise w with standard deviation σ is added to the output.

Based on the input observability Gramian $\mathcal{O}_{U_M,M}$, $M = 0, 1, \dots$ and (13), the required standard deviation σ to achieve (ε, δ)-differential private at each M is computed for $b = 0.1$, $\varepsilon = 0.1$ and $\delta = 0.1$ and is shown in Fig.2. The required standard deviations increase as the duration increases, since more data one collects, less private a system becomes. Therefore, for a long duration, one needs to add large noise. However, output y^w with large noise may not be helpful for data analysis. An ad hoc idea addressing this problem is changing the standard deviation of noise at each M based on $\mathcal{O}_{U_M,M}$, and studying differential privacy with time varying deviation is a topic for our future work.

V. Conclusion

In this paper, we have clarified that attack detection and privacy preservation can be analyzed in the same input observability framework. To detect attacks, we constructed an input observer for an input observable system. As a measure of privacy, we extended the concept of Gramian to input observability and then showed that differential privacy can be evaluated by the maximum eigenvalue of the input observability Gramian. We are currently working on other forms of attacks and concepts of privacy. We are also interested in studying nonlinear dynamic processes, and some preliminary results have been summarized in [23].

References

[1] P. McDaniel and S. McLaughlin, "Security and privacy challenges in the smart grid," *IEEE Security & Privacy*, vol. 7, no. 3, 2009.

[2] S. R. Islam, D. Kwak, M. H. Kabir, M. Hossain, and K.-S. Kwak, "The internet of things for health care: a comprehensive survey," *IEEE Access*, vol. 3, pp. 678–708, 2015.

[3] NCCIC/ICS-CERT, "Cyber-attack against ukrainian critical infrastructure," https://ics-cert.us-cert.gov/alerts/IR-ALERT-H-16-056-01.

[4] A. Greenberg, "How one of Apple's key privacy safeguards falls short," https://www.wired.com/story/apple-differential-privacy-shortcomings/, 2017.

[5] J. Tang, A. Korolova, X. Bai, X. Wang, and X. Wang, "Privacy loss in Apple's implementation of differential privacy on macOS 10.12," *arXiv preprint arXiv:1709.02753*, 2017.

[6] N. Ye, J. Giordano, and J. Feldman, "A process control approach to cyber attack detection," *Communications of the ACM*, vol. 44, no. 8, pp. 76–82, 2001.

[7] L. Feinstein, D. Schnackenberg, R. Balupari, and D. Kindred, "Statistical approaches to DDoS attack detection and response," *Proceedings of the DARPA Information Survivability Conference and Exposition*, pp. 303–314, 2003.

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

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

[10] A. Teixeira, H. Sandberg, and K. H. Johansson, "Networked control systems under cyber attacks with applications to power networks," *Proceedings of the 2010 American Control Conference*, pp. 3690–3696, 2010.

[11] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," *Proceedings of the 3rd Theory of Cryptography Conference*, pp. 265–284, 2006.

[12] J. L. Ny and G. Pappas, "Differentially private filtering," *IEEE Transactions on Automatic Control*, vol. 59, no. 2, pp. 341–354, 2014.

[13] J. Cortés, G. E. Dullerud, S. Han, J. Le Ny, S. Mitra, and G. J. Pappas, "Differential privacy in control and network systems," *Proceedings of the 55th IEEE Conference on Decision and Control*, pp. 4252–4272, 2016.

[14] R. Anguluri, R. Dhal, S. Roy, and F. Pasqualetti, "Network invariants for optimal input detection," *Proceedings of the 2016 American Control Conference*, pp. 3776–3781, 2016.

[15] J. L. Massey and M. K. Sain, "Inverses of linear sequential circuits," *IEEE Transactions on Computers*, vol. 17, no. 4, pp. 330–337, 1968.

[16] M. K. Sain and J. L. Massey, "Invertibility of linear time-invariant dynamical systems," *IEEE Transactions on Automatic Control*, vol. 14, no. 2, pp. 141–149, 1969.

[17] P. J. Moylan, "Stable inversion of linear systems," *IEEE Transactions on Automatic Control*, vol. 22, no. 1, pp. 74–78, 1977.

[18] M. Hou and R. J. Patton, "Input observability and input reconstruction," *Automatica*, vol. 34, no. 6, pp. 789–794, 1998.

[19] P. Sannuti and A. Saberi, "Special coordinate basis for multivariable linear systems – finite and infinite zero structure, squaring down and decoupling," *International Journal of Control*, vol. 45, no. 5, pp. 1655–1704, 1987.

[20] S. Lang, *Algebra*. New York: Springer-Verlag, 2002.

[21] R. E. Kalman, "Contributions to the theory of optimal control," *Boletín de la Sociedad Matemática Mexicana*, vol. 5, no. 2, pp. 102–119, 1960.

[22] B. Moore, "Principal component analysis in linear systems: Controllability, observability, and model reduction," *IEEE Transactions on Automatic Control*, vol. 26, no. 1, pp. 17–32, 1981.

[23] Y. Kawano and M. Cao, "Differential privacy and qualitative privacy analysis for nonlinear dynamical systems," *Proceedings of the 7th IFAC Workshop on Distributed Estimation and Control in Networked Systems*, pp. 52–57, 2018.