Low Complexity Secure State Estimation Design for Linear System with Non-derogatory A

Zishuo Li and Yilin Mo

Abstract— **We consider the problem of estimating the state of a time-invariant linear Gaussian system in the presence of integrity attacks. The attacker can compromise p out of m sensors, the set of which is fixed and unknown to the system operator, and manipulate the measurements arbitrarily. Under the assumption that all the eigenvalues of system matrix A have geometric multiplicity 1 (A is non-derogatory), we propose a secure estimation scheme that is resilient to integrity attack as long as the system is $2p$-sparse observable. In the absence of attack, the proposed estimation coincides with Kalman estimation with a certain probability that can be adjusted. Furthermore, our proposed estimator is computational efficient during the security condition checking in the designing phase and during the estimation computing in the online operating phase. A numerical example is provided to corroborate the results and illustrate the performance of the proposed estimator.**

I. Introduction

Cyber-Physical System (CPS) and Internet of Things (IoT) are playing an increasingly important role in critical infrastructures and everyday life. The related area of CPS and IoT continues to emerge and expand as costs drop and the confluence of sensors, platforms and networks increases [1]. Simultaneously, the cyber-security risks and attack surfaces are also increasing [2] since CPS relies on remote sensing devices, communication channels, and spatially distributed processors, which are prone to failures under cyber attacks on the data acquisition and communication channels. The unintentional faults or malicious attacks could cause severe system damage, economic loss, and environmental degradation, e.g., the Stuxnet launched on Iran's nuclear facilities [3], power blackouts in Brazil [4] and Ukraine [5], etc. The research community has recognized the importance of CPS security, especially the design of secure detection, estimation, and control strategy [6].

Recently, substantial research efforts have been devoted to secure state estimation against malicious sensors. One of the research paths is the finite horizon approach, where only measurements in a finite time-window are considered when recovering the system state. The problem is combinatorial in nature since searching for the corrupted sensors involves minimizing the ℓ_0 norm [7], which is an NP problem. Fawzi et al. [7] propose an algorithm using its convex relaxation, i.e., ℓ_1 optimization, to solve the problem in the absence of noise. Similarly, Shoukry and Tabuada [8] adopt a 2-norm batch optimization approach to solve the state estimation problem. Shoukry et al. [9] propose an algorithm in the presence of noise by searching for reliable sensors using consistency check, and the searching complexity is reduced based on Satisfiability Modulo Theory. However, in these works, the sensory data out of the time window are discarded, which may cause performance degradation and estimation delay.

Another solution is the switch estimator [10] [11] where multiple estimates are maintained based on measurements from different subsets of sensors, and the system operator switches between these estimates based on the evaluation of their reliability by consistency checking or malicious detection algorithms. However, the combinatorial nature of estimation candidates or sensor subsets may incur heavy computation and storage burden of the devices. In view of these problems, Liu et al. [12] design local estimators whose weighted sum coincides with the Kalman estimate with a certain probability in the absence of attack. The local estimates are fused securely by a quadratic programming problem with an ℓ_1 term to handle the sparse outliers.

Even though there are efficient algorithms computing the estimation by introducing ℓ_1 relaxation, the design of the secure state estimation is still computational challenging. For secure state recovery problem in the presence of p malicious sensors, it is required that the system needs to be $2p$-sparse observable [8], and calculating the sparse observability index for general systems is proved to be NP-hard [13]. Besides sparse observability, other results [7] [12] [14] impose stronger conditions for system resiliency, whose validations are also NP problems. However, it has been observed by Mao et al. [13] that when the eigenvalues of the system matrix A have unitary multiplicity (A is non-derogatory), the computational complexity of secure state reconstruction without noise can be significantly reduced. Leveraging upon this assumption, we propose our secure estimation scheme in the formulation of LTI system with Gaussian noise, and it has the following merits:

- In the presence of p compromised sensors, the proposed estimation is secure if the system is $2p$-sparse observable. In the absence of attack, the proposed estimation coincides with Kalman estimation with certain probability, which can be adjusted.
- During the designing phase, the sparse observability index can be computed with low complexity. During the algorithm operating phase, the proposed estimation is formulated as the solution of a convex optimization problem which can be computed efficiently.

Zishuo Li and Yilin Mo are with the Department of Automation and BNRist, Tsinghua University, Beijing, China. Email: {lizs19@mails.tsinghua.edu.cn, ylmo@mail.tsinghua.edu.cn}. This work is supported by the National Key Research and Development Program of China under Grant 2018AAA0101601.

Organization: We introduce the problem formulation and preliminary results in Section II. The main results are provided in Section III and collaborated by numerical simulation in Section IV. Section V finally concludes the paper.

Notations: Cardinality of a set $\mathcal{S}$ is denoted as $|\mathcal{S}|$. Conjugate transpose of a matrix A is denoted as A'. The determinant of a matrix is represented by $\det(\cdot)$. Diagonal matrix with diagonal elements $A_1, \cdots, A_k$ is denoted as $\operatorname{diag}(A_1, \cdots, A_k)$. Denote the span of row vectors of matrix A as $\operatorname{rowspan}(A)$. All-one vector with size $m \times 1$ is denoted as $\mathbf{1}_m$. I_n is the identity matrix with size $n \times n$. The i-th entry of a vector x is represented by x_i or $[x]_i$. $\|\cdot\|_\infty$ represents the infinity vector norm or (induced) infinity matrix norm which is clear according to the context.

II. Problem Formulation and Preliminary Results

A. Secure dynamic state estimation

In this paper, we consider the linear time-invariant system with Gaussian noise:

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

where $x(k) \in \mathbb{R}^n$ is the system state, $w(k) \sim N(0, Q)$ and $v(k) \sim N(0, R)$ are i.i.d. Gaussian process noise and measurement noise with zero mean and covariance matrix Q and R. The vector $y(k) \in \mathbb{R}^m$ is the collection of measurement from all m sensors, and i-th entry $y_i(k)$ is the measurement from sensor i. The vector $a(k)$ denotes the bias injected by an adversary and $a_i(k)$ is the attack on sensor i. Define

$$z(k) = Cx(k) + v(k)$$

as the true measurements without the attack. The initial state $x(0) \sim N(0, \Sigma)$ is assumed to be zero mean Gaussian and is independent from the process noise $\{w(k)\}$.

The secure dynamic estimation problem aims at recovering system state $x(k)$ at every time k based on all history observations $\{y(t), 0 \le t \le k\}$ which have been partly manipulated by the malicious attacker. It is conventional in the literature [7] [9] that the attacker can only compromise a fixed subset of sensors with known maximum cardinality. Denote the index set of all sensors as $\mathcal{R} \triangleq \{1, 2, \ldots, m\}$. For any index set $\mathcal{I} \subseteq \mathcal{R}$, define the complement set to be $\mathcal{I}^c \triangleq \mathcal{R} \backslash \mathcal{I}$. Define the support of vector $a \in \mathbb{R}^n$ as $\operatorname{supp}(a) \triangleq \{i | 1 \le i \le n, a_i \neq 0\}$ where a_i is the i-th entry of vector a. We have the following assumptions on the malicious adversary.

Definition 1 (Sparse Attack) The attack is called a (p, m)-sparse attack if the vector sequence $a(k)$ satisfy that, there exists a time invariant index set $\mathcal{I} \subseteq \mathcal{R}$ with $|\mathcal{I}| = p$ such that $\bigcup_{k=1}^{\infty} \operatorname{supp}\{a(k)\} = \mathcal{I}$.

Closely related to the sparse attack, we introduce the notion of sparse observability that characterizes the system observability in the presence of attack.

Definition 2 (Sparse observability) The sparse observability index of system (1)-(2) is the largest integer s such that system[1] $(A, C_{\mathcal{R}\backslash\mathcal{I}})$ is observable for any set of sensors $\mathcal{I} \subset \mathcal{R}$ with cardinality $|\mathcal{I}| = s$. When the sparse observability index is s, we say that the system with pair (A, C) is s-sparse observable.

Define $y(k_1 : k_2)$ as the sequence $\{y(k_1), y(k_1+1), \cdots, y(k_2)\}$. Similar notation is also applied on $z(k)$. For linear Gaussian noise system, the estimation is secure if the estimation error is bounded by a constant term irrelevant to the attack.

Definition 3 (Secure estimator) An estimator is an infinite sequence of mappings $g = \{g_k\}_{k=1}^{\infty}$ where g_k is a mapping from all the history observations to an estimation at time k:

$$g_k(y(0:k)) = \hat{x}(k).$$

Define the estimation difference introduced by attack as

$$q_k \triangleq \|g_k(z(0:k)) - g_k(y(0:k))\|_2 .$$

The estimator is said to be secure against the (p, m)-sparse attack if the following holds:

$$\sup_{k \in \mathbb{Z}^+} \mathbb{E}\left[q_k^2\right] < \infty,$$

where $\mathbb{E}$ is the expectation with respect to the probability measure generated by the Gaussian noise $\{w(k)\}$ and $\{v(k)\}$.

If all sensors are benign, i.e., $a(k) = \mathbf{0}$ for all k, the optimal state estimator is the classical Kalman filter:

$$\begin{aligned} \hat{x}(k) &= \hat{x}(k|k-1) + K(k)\left[y(k) - C\hat{x}(k|k-1)\right], \\ P(k) &= P(k|k-1) - K(k)CP(k|k-1), \end{aligned}$$

where

$$\begin{aligned} &\hat{x}(k|k-1) = A\hat{x}(k-1), P(k|k-1) = AP(k-1)A' + Q, \\ &K(k) = P(k|k-1)C'\left(CP(k|k-1)C' + R\right)^{-1}, \end{aligned}$$

with initial condition $\hat{x}(0|-1) = 0,\ P(0|-1) = \Sigma$. It is well-known that for observable system, the estimation error covariance matrices $P(k)$ and the gain $K(k)$ will converge to

$$\begin{aligned} P &\triangleq \lim_{k \to \infty} P(k),\ P_+ = APA' + Q, \\ K &\triangleq P_+ C'\left(CP_+C' + R\right)^{-1}. \end{aligned}$$

Since typically the control system will be running for an extended period of time, we focus on the case where the Kalman filter is in steady state, and thus the Kalman filter reduces to the following fixed-gain linear estimator:

$$\hat{x}(k+1) = (A - KCA)\hat{x}(k) + Ky(k+1). \tag{3}$$

Before introducing our work, we first recall some results in the previous work that decomposes the fix gain Kalman filter to local estimates and recovers it securely by an optimization problem.

[1] The matrix $C_{\mathcal{R}\backslash\mathcal{I}}$ represents the matrix composed of rows of C with row index in $\mathcal{R}\backslash\mathcal{I}$.

B. Preliminary Results

The following preliminary results in this subsection are from [12]. We introduce the following assumption:

Assumption 1 $A-KCA$ has n distinct eigenvalues. Moreover, $A-KCA$ and A do not share any eigenvalue.

Since $A-KCA$ has distinct eigenvalues, it can be diagonalized as:

$$A-KCA=V\Pi V^{-1}. \tag{4}$$

Define the eigenvalues of $A-KCA$ as $\pi_1,\cdots,\pi_n$. Consider local estimation $\zeta_i(k)$ which is the system response of sensor i. The local estimator satisfies the following dynamic:

$$\zeta_i(k+1)=\Pi\zeta_i(k)+\mathbf{1}_n y_i(k+1). \tag{5}$$

Define G_i as

$$G_i \triangleq \begin{bmatrix} C_iA\left(A-\pi_1 I\right)^{-1} \\ \vdots \\ C_iA\left(A-\pi_n I\right)^{-1} \end{bmatrix}. \tag{6}$$

It has been proved in [12] Corollary 1 that $\zeta_i(k)$ is a stable estimation of $G_ix(k)$ and the difference between them is characterized in the follow lemma.

Lemma 1 ([12]) Let $\varepsilon_i(k)\triangleq\zeta_i(k)-G_ix(k)$, then

$$\begin{aligned}\varepsilon_i(k+1)=\Pi\varepsilon_i(k)&+\left(G_i-\mathbf{1}_nC_i\right)w(k)\\ &-\mathbf{1}_nv_i(k+1)-\mathbf{1}_na_i(k+1).\end{aligned} \tag{7}$$

Define $\tilde{Q}\in\mathbb{C}^{mn\times mn}$ as the covariance of noise term $\left(G_i-\mathbf{1}_nC_i\right)w(k)-\mathbf{1}_nv_i(k+1)$ for all i, i.e.,

$$\tilde{Q}\triangleq\begin{bmatrix}G_1-\mathbf{1}_nC_1\\ \vdots\\ G_m-\mathbf{1}_nC_m\end{bmatrix}Q\begin{bmatrix}G_1-\mathbf{1}_nC_1\\ \vdots\\ G_m-\mathbf{1}_nC_m\end{bmatrix}'+R\otimes\mathbf{1}_{n\times n}, \tag{8}$$

where $\otimes$ is the Kronecker product. Define $\tilde{\Pi}\in\mathbb{C}^{mn\times mn}$ as

$$\tilde{\Pi}\triangleq\begin{bmatrix}\Pi & & \\ & \ddots & \\ & & \Pi\end{bmatrix}.$$

The stable covariance of $\varepsilon(k)\triangleq[\varepsilon_1(k)',\cdots,\varepsilon_m(k)']'$ is the solution $\tilde{W}$ of the following Lyapunov equation:

$$\tilde{W}=\tilde{\Pi}\tilde{W}\tilde{\Pi}'+\tilde{Q}.$$

The matrix $\tilde{W}$ is well-defined since Π is strictly stable. As a result, the secure estimation can be recovered by the solution of the following optimization problem where $\zeta(k)\triangleq[\zeta_1(k)',\cdots,\zeta_m(k)']'$ and $G\triangleq[G'_1,\cdots,G'_m]'$.

$$\underset{\check{x}(k),\mu(k),\nu(k)}{\text{minimize}}\quad \frac{1}{2}\mu(k)'\tilde{W}^{-1}\mu(k)+\gamma\|\nu(k)\|_1 \tag{9a}$$

$$\text{subject to}\quad \zeta(k)=G\check{x}(k)+\mu(k)+\nu(k). \tag{9b}$$

The parameter γ is a non-negative constant chosen by the system operator. According to [12], the solution $\check{x}(k)$ to problem (9) is a secure estimation and has the following properties.

Theorem 1 ([12] Theorem 4) In the presence of (p,m)-sparse attack, the state estimation $\check{x}(k)$ is secure if the following inequality holds for all $u\neq\mathbf{0}$, $u\in\mathbb{R}^n$:

$$\sum_{i\in\mathcal{I}}\|G_iu\|_1<\sum_{i\in\mathcal{I}^c}\|G_iu\|_1,\quad\forall\,\mathcal{I}\subset\mathcal{R},|\mathcal{I}|\le p. \tag{10}$$

Even though Theorem 1 establishes the sufficient condition of the estimation to be secure, validating (10) is NP-hard. In the following section, we reduce the complexity of checking condition (10) by transforming G_i to its canonical form under the assumption that the geometric multiplicities of all the eigenvalues of A are 1.

III. Main Results

In this section, under the assumption on A, we first prove that the span of rows of G_i coincides with the observability space of (A,C_i), which implies that the matrix G_i has a canonical form under row operations. Based on the canonical form, a new secure estimation scheme is proposed.

A. Canonical form of G_i

In order to prevent degeneration problem, we introduce the following assumption.

Assumption 2 System matrix A is non-singular, and all the eigenvalues of A have geometric multiplicity 1. Without loss of generality, we assume that A is in the following Jordan canonical form:

$$A=\begin{bmatrix}J_1&\mathbf{0}&\cdots&\mathbf{0}\\ \mathbf{0}&J_2&\cdots&\mathbf{0}\\ \vdots&\vdots&\ddots&\vdots\\ \mathbf{0}&\mathbf{0}&\cdots&J_l\end{bmatrix},\ J_k=\begin{bmatrix}\lambda_k&1&0&\cdots&0\\ 0&\lambda_k&1&\cdots&0\\ \vdots&\ddots&\ddots&\ddots&\vdots\\ \vdots&\ddots&\ddots&\ddots&1\\ 0&\cdots&\cdots&0&\lambda_k\end{bmatrix},$$

where $\lambda_l\neq\lambda_k$ when $l\neq k$.

Define the observable matrix of system (A,C_i) as

$$O_i\triangleq\left[\ C_i'\ \middle|\ (C_iA)'\ \middle|\ \cdots\ \middle|\ \left(C_iA^{n-1}\right)'\ \right]'. \tag{11}$$

Before continuing on, we need the following notation of state observability. Define $\mathcal{S}_j$ as the index set of sensors that can observe state j, i.e.

$$\mathcal{S}_j\triangleq\{i\in\mathcal{R}\mid O_i'e_j\neq\mathbf{0}\}, \tag{12}$$

where $\mathcal{R}\triangleq\{1,2,\cdots,m\}$ is the index set of all sensors and e_j is the canonical basis vector with 1 on the j-th entry and 0 on the other entries. We have the following theorem characterizing the structure of G_i. The proof is provided in Appendix A.

Theorem 2 Assume system matrix A satisfies Assumption 2, then the following equation holds:

$$\operatorname{rowspan}(G_i)=\operatorname{rowspan}(O_i)=\operatorname{rowspan}(H_i), \tag{13}$$

where H_i is the following diagonal matrix

$$H_i \triangleq \begin{bmatrix} \mathbb{I}_{i\in\mathcal{S}_1} & & & \\ & \mathbb{I}_{i\in\mathcal{S}_2} & & \\ & & \ddots & \\ & & & \mathbb{I}_{i\in\mathcal{S}_n} \end{bmatrix},$$

and $\mathbb{I}_{\mathscr{E}}$ is the indicator function that takes the value 1 when $\mathscr{E}$ is true and value 0 when $\mathscr{E}$ is not.

According to Theorem 2, one directly obtains the following corollary:

Corollary 1 For every sensor index $i \in \mathcal{R}$, there exists an invertible matrix P_i such that $P_iG_i = H_i$.

After transformation P_i, matrix G_i is transformed into canonical form H_i whose rows are either canonical basis vectors or zero vectors. The non-zero entries of H_i records the state observability of sensor i. Therefore, the sparse observability index can be directly obtained from H_i.

Corollary 2 Denote s as the sparse observability index of system (1)-(2). Then

$$s = \min_{j\in\{1,2,\cdots,n\}} |\mathcal{S}_j| - 1.$$

Proof: For arbitrary $\bar{s}$ that satisfy $\bar{s} \geq s+1$, there exists a state index j^* and a sensor index set $\mathcal{I}^*$ with $|\mathcal{I}^*| = \bar{s}$ such that $\mathcal{S}_{j^*} \cap (\mathcal{R}\setminus\mathcal{I}^*) = \varnothing$. As a result, state j^* can not be observed by any sensor in $\mathcal{R}\setminus\mathcal{I}^*$, i.e.,

$$e_{j^*} \notin \mathrm{rowspan}(O_i),\ \forall i \in \mathcal{R}\setminus\mathcal{I}^*.$$

and thus system $(A, C_{\mathcal{R}\setminus\mathcal{I}^*})$ is not observable. For arbitrary $\underline{s}$ that satisfy $\underline{s} \leq s$, arbitrary j and arbitrary $\mathcal{I}$ with $|\mathcal{I}| = \underline{s}$, one obtains $\mathcal{S}_{j^*} \cap (\mathcal{R}\setminus\mathcal{I}^*) \neq \varnothing$, which means for all j, there exists $i^* \in \mathcal{R}\setminus\mathcal{I}$ such that: $e_j \in \mathrm{rowspan}(O_{i^*})$. Therefore, system $(A, C_{\mathcal{R}\setminus\mathcal{I}})$ is observable. According to Definition 2, the system is s-sparse observable. ∎

In summary, under Assumption 2, the system sparse observability index can be obtained with low computation complexity. In the following subsection, we will propose an information fusion scheme that is secure in the presence of (p,m)-sparse attack as long as the system is $2p$-sparse observable.

B. Secure Information Fusion

Recalling the transformation P_i introduced in Corollary 1, define $\tilde{P} \triangleq \mathrm{diag}(P_1,\cdots,P_m)$, $\tilde{M} \triangleq \tilde{P}\tilde{W}\tilde{P}'$ and

$$\mathcal{Y}(k) \triangleq \begin{bmatrix} P_1\zeta_1(k) \\ \vdots \\ P_m\zeta_m(k) \end{bmatrix},\ H \triangleq \begin{bmatrix} H_1 \\ \vdots \\ H_m \end{bmatrix}. \tag{14}$$

We present the following optimization problem whose solution is our proposed estimation. The constant γ is an adjustable parameter.

$$\underset{\tilde{x}(k),\mu(k),\nu(k)}{\text{minimize}} \quad \frac{1}{2}\mu(k)'\tilde{M}^{-1}\mu(k) + \gamma\|\nu(k)\|_1 \tag{15a}$$

$$\text{subject to} \quad \mathcal{Y}(k) = H\tilde{x}(k) + \mu(k) + \nu(k). \tag{15b}$$

We have the following theorem demonstrating that the estimation is secure under attack.

Theorem 3 In the presence of arbitrary (p,m)-sparse attack, if the system (A,C) is $2p$-sparse observable, the estimation $\tilde{x}(k)$ solved from (15) is secure.

The proof of Theorem 3 is provided in Appendix B. Since sparse observability index only requires simple computation according to Corollary 2, this work reduces the complexity of evaluating system vulnerability significantly under the assumption of geometric multiplicity. For general A that has eigenvalues with geometric multiplicity larger than 1 (A is derogatory), computing sparse observability is an NP-hard problem [13], and there is no computational efficient solution unless P=NP. Simultaneously, for algorithm online operation, the computing of estimation involves solving a convex optimization problem based on LASSO [15], which can be done efficiently. Moreover, our proposed estimator holds optimality in the absence of attack for certain probability, which can be adjusted by parameter γ.

IV. Illustrative Example

We use an inverted pendulum for the numerical simulation[2]. The physical parameters are illustrated in Fig. 1. The control input $u(k)$ is the force applied on the cart, and we assume that there is no frictions of any form. The state x_1, x_2, x_3, x_4 represent cart position coordinate, cart velocity, pendulum angle from vertical and pendulum angle velocity respectively.

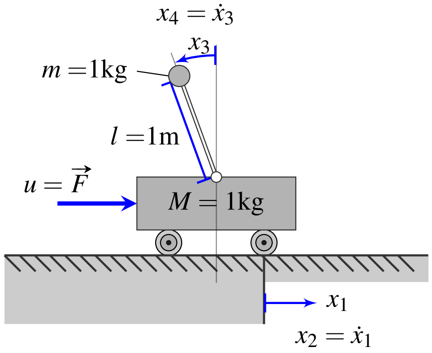

Fig. 1. Illustration of the inverted pendulum.

Consider the system linearized at $x_3 = x_4 = 0$, and we sample the continuous-time linear system periodically every 0.1 seconds. The system matrix A can be transformed into the following Jordan canonical form

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0.642 & 0 \\ 0 & 0 & 0 & 1.557 \end{bmatrix},$$

and there is a Jordan block with size 2 × 2 and the geometric multiplicity of all the eigenvalues are 1.

The output matrix C satisfy $C_1 = C_2 = C_3 = [1\ \ 0\ \ 0\ \ 0]$, $C_4 = [0\ \ 0\ \ 1\ \ 0]$, i.e., first three sensors are monitoring the

[2] The corresponding code is posted on `https://github.com/zs-li/resilient_dynamic_estimation`.

cart position x_1 and the last sensor is monitoring pendulum angle x_3. The noise covariances of the system satisfy $w(k) \sim N(0, 0.001 \times I_4)$ and $v(k) \sim N(0, 0.001 \times I_4)$. According to Corollary 2, the system is 2-sparse observable and our proposed estimator secure in the presence of 1 corrupted sensor. The controller of the system is designed as a Linear-Quadratic Regulator (LQR), and the feedback matrix is chosen as

$$K_{\mathrm{lqr}} = \begin{bmatrix} -0.604 & -1.678 & -39.514 & -9.721 \end{bmatrix}.$$

We demonstrate the performance of estimation on close-loop system where $u(k) = -K_{\mathrm{lqr}} x(k)$. Fig. 2 presents the performance of the estimation of system state x_3 (pendulum angle). In the absence of attack, our proposed estimation substantially coincides with the Kalman estimation. The numerical difference attributes to large Gaussian noise that occurs occasionally and accuracy loss in numerical calculation. The bottom subplot represents the estimation under attack on sensor 4. The attack $a_4(k)$ is a time-independent random value uniformly distributed on interval $(-5, 5)$. As shown in the bottom of Fig 2, Kalman estimation (denoted as red dashed line) has larger estimation error than our proposed estimation under the attack.

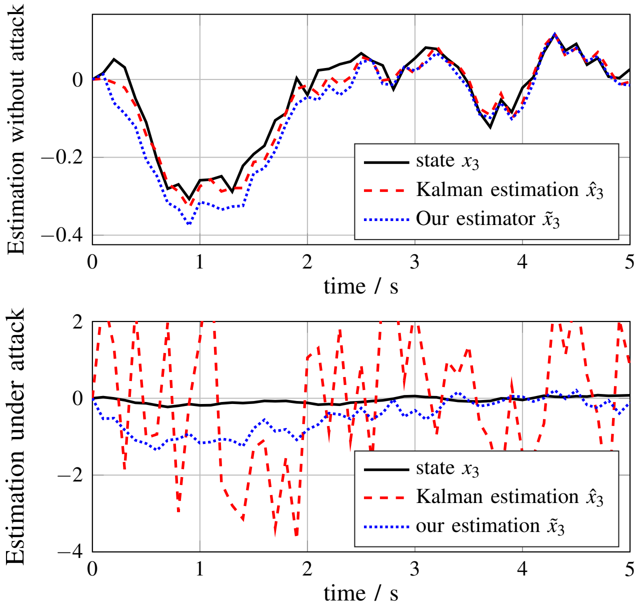

Fig. 2. Estimation of state x_3 (pendulum angle) in the absence and in the presence of attack. The initial state is $x(0) = [0,\ 1,\ 0,\ 1]'$.

Fig. 3 illustrates the estimation mean square error (MSE) of our proposed estimator with varying γ in the absence and in the presence of attack on sensor 4. The attack $a_4(k)$ is the same as in Fig. 2. The MSE of the oracle Kalman estimation, .i.e., Kalman estimation not affected by the attack, is illustrated by the red dashed line in Fig. 3. As shown in the figure, by properly choosing γ, the MSE of our proposed estimator is smaller than that of Kalman estimation (the red horizontal line), with the cost that MSE without attack is slightly larger.

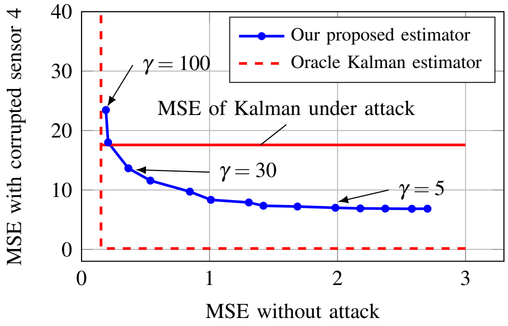

Fig. 3. Estimation mean square error (MSE) in the absence and in the presence of attack with varying tuning parameter γ.

V. Conclusion

This paper considers LTI system with Gaussian noise against malicious attack on a subset of sensors. We improve upon the previous work by reducing the computation complexity in the designing phase while remaining low computation complexity for online estimation, under the assumption that all the geometric multiplicities of eigenvalues of A are 1 (A is non-derogatory). To achieve this, we prove that the span of the rows of G_i is equivalent to the observable space, based on which the canonical form H_i is designed. The proposed estimator is formulated as a convex optimization problem based on H_i. We further prove that the proposed estimation is secure if the system is $2p$-sparse observable, which is easily checked due to the simple form of H_i. Moreover, in the absence of attack, the proposed estimation coincides with Kalman estimation for certain probability, which can be adjusted by tuning parameter γ to balance between the performance with and without attack.

Appendix

A. Proof of Theorem 2

Due to space limit, we provide a sketch of proof.
Proof: (**Skecth of proof**) Define the characteristic polynomial of A as $p(x) = a_n x^n + \cdots + a_1 x + a_0$. Define polynomial fraction $q_\pi(x)$ with respect to constant π as $q_\pi(x) = \frac{p(x) - p(\pi)}{x - \pi}$ where $x \neq \pi$. Therefore,

$$q_\pi(A)(A - \pi I) = p(A) - p(\pi)I = -p(\pi)I,$$

where the last equality comes from Cayley-Hamilton Theorem. As a result, when π is not the eigenvalue of A,

$$(A - \pi I)^{-1} = -\frac{1}{p(\pi)} q_\pi(A) \tag{16}$$

In order to simplify notations, we define

$$b_{j,k} \triangleq -\frac{1}{p(\pi_j)} \sum_{i=0}^{n-k-1} a_{i+k+1} \pi_j^i, \tag{17}$$

where π_j is the j-th diagonal element of Π, i.e., j-th eigenvalue of $A - KCA$ as defined in (4). According to (16), the j-th row of matrix G_i can be reformulated as

$$C_i A (A - \pi_j I)^{-1} = \begin{bmatrix} b_{j,0} & b_{j,1} & \cdots & b_{j,n-1} \end{bmatrix} O_i A.$$

Therefore, G_i can be interpreted as $G_i = TO_iA$ where T is an invertible matrix, and thus $\text{rowspan}(G_i) = \text{rowspan}(O_iA)$. Since A is assumed to be invertible, one can proof that $\text{rowspan}(O_i) = \text{rowspan}(O_iA)$. According to Assumption 2, nonzero columns of O_i are linear independent, and equivalently nonzero columns of G_i are linear independent. Therefore, $i \in \mathcal{S}_j$ is equivalent to that j-th column of G_i is non-zero, i.e., G_i has the same row-span with the canonical form H_i. ■

B. Proof of Theorem 3

Before proving Theorem 3, we need the following Lemma. Define the number of honest sensors and compromised sensors (w.r.t. compromised set $\mathcal{I}$) that can observe state j as:

$$h_j(\mathcal{I}) \triangleq |\mathcal{S}_j \cap \mathcal{I}^c|, \; c_j(\mathcal{I}) \triangleq |\mathcal{S}_j \cap \mathcal{I}|.$$

We have the following lemma quantifying the property of $h_j(\mathcal{I})$ and $c_j(\mathcal{I})$.

Lemma 2 The following two propositions are equivalent.

(1) The system is $2p$-sparse observable.

(2) For any $\mathcal{I}$ with $|\mathcal{I}| = p$, the inequality $c_j(\mathcal{I}) < h_j(\mathcal{I})$ holds for all $j \in \{1, 2, \cdots, n\}$.

Proof: (**Proof of Lemma 2**) We prove the contrapositive of (1)⇒(2). Supposing that there exists j^* and $\mathcal{I}^*$ with $|\mathcal{I}^*| = p$ such that $c_{j^*}(\mathcal{I}^*) \geq h_{j^*}(\mathcal{I}^*)$, then $h_{j^*}(\mathcal{I}^*) \leq c_{j^*}(\mathcal{I}^*) \leq |\mathcal{I}^*| = p$. Noticing that $c_j(\mathcal{I}) + h_j(\mathcal{I}) = |\mathcal{S}_j|$ holds for all $\mathcal{I}$, we have $|\mathcal{S}_{j^*}| \leq 2p$. There exists set $\mathcal{A}$ that satisfy $\mathcal{A} \supseteq \mathcal{S}_{j^*}$ and $|\mathcal{A}| = 2p$. According to the definition of $\mathcal{S}_{j^*}$, there exists no sensor in set $\mathcal{R} \setminus \mathcal{A}$ who can observe state j^*, i.e.,

$$e_{j^*} \notin \text{rowspan}(O_i), \; \forall i \in \mathcal{R} \setminus \mathcal{A}.$$

As a result, system $(A, C_{\mathcal{R} \setminus \mathcal{A}})$ is not $2p$-sparse observable according to Definition 2.

We proceed to prove (2)⇒(1). Since for any $\mathcal{I}$ with $|\mathcal{I}| = p$, $h_j(\mathcal{I}) > c_j(\mathcal{I}) \geq 0$, the system sparse observability index is at least p. Therefore, for each $j \in \{1, 2, \cdots, n\}$, there exists an $\mathcal{I}^*$ such that $c_j(\mathcal{I}^*) = p$, and thus $|\mathcal{S}_j| = h_j(\mathcal{I}^*) + c_j(\mathcal{I}^*) \geq 2p + 1$. According to the definition $\mathcal{S}_j$, there are at least $2p+1$ sensors that can observe sensor j, and the system is $2p$-sparse observable. ■

Proof: (**Proof of Theorem 3**) In view of Theorem 1 and Lemma 2, it suffices to prove that the following two propositions are equivalent:

(1) The system is $2p$-sparse observable.

(2) The following inequality holds for all $x \neq \mathbf{0}$, $x \in \mathbb{R}^n$:

$$\sum_{i \in \mathcal{I}} \|H_i x\|_1 < \sum_{i \in \mathcal{I}^c} \|H_i x\|_1, \quad \forall \, \mathcal{I} \subset \mathcal{R}, |\mathcal{I}| \leq p. \tag{18}$$

Based on the form of H_i in Theorem 2, inequality (18) can be written as

$$\sum_{j=1}^{n} \sum_{i \in \mathcal{I} \cap \mathcal{S}_j} |x_j| < \sum_{j=1}^{n} \sum_{i \in \mathcal{I}^c \cap \mathcal{S}_j} |x_j|$$

or equivalently

$$\sum_{j=1}^{n} c_j(\mathcal{I}) \cdot |x_j| < \sum_{j=1}^{n} h_j(\mathcal{I}) \cdot |x_j|.$$

If the system is $2p$-sparse observable, we have $c_j(\mathcal{I}) < h_j(\mathcal{I})$ for all $j \in \{1, 2, \cdots, n\}$. Thus, $\sum_{j=1}^{n} (h_j(\mathcal{I}) - c_j(\mathcal{I})) \cdot |x_j| > 0$ holds for all $x \neq \mathbf{0}$. If the system is not $2p$-sparse observable, there exists $\mathcal{I}^*$ with $|\mathcal{I}^*| = p$ and j^* such that $c_{j^*}(\mathcal{I}^*) > h_{j^*}(\mathcal{I}^*)$. One can design $x \neq \mathbf{0}$ such that

$$(c_{j^*}(\mathcal{I}^*) - h_{j^*}(\mathcal{I}^*)) \cdot |x_{j^*}| > \sum_{j \neq j^*} (h_j(\mathcal{I}^*) - c_j(\mathcal{I}^*)) \cdot |x_j|.$$

Therefore, condition (18) is violated and the proof is completed. ■

References

[1] U. D. of Homeland Security (DHS) S and T. C. S. Division, "Cyber security division technology guide 2018," 2018. [Online]. Available: https://www.dhs.gov/publication/st-technology-guide

[2] A. A. Cárdenas, S. Amin, and S. Sastry, "Research challenges for the security of control systems." *HotSec*, vol. 5, p. 15, 2008.

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

[4] J. Conti, "The day the samba stopped [power blackouts]," *Engineering and Technology*, vol. 5, 03 2010.

[5] G. Liang, S. R. Weller, J. Zhao, F. Luo, and Z. Y. Dong, "The 2015 ukraine blackout: Implications for false data injection attacks," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 3317–3318.

[6] A. Cardenas, S. Amin, B. Sinopoli, A. Giani, A. Perrig, S. Sastry *et al.*, "Challenges for securing cyber physical systems," in *Workshop on future directions in cyber-physical systems security*, vol. 5, no. 1. Citeseer, 2009.

[7] 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.

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

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

[10] Y. Nakahira and Y. Mo, "Attack-resilient $\mathcal{H}_2$, $\mathcal{H}_\infty$, and ℓ_1 state estimator," *IEEE Transactions on Automatic Control*, vol. 63, no. 12, pp. 4353–4360, 2018.

[11] A. Lu and G. Yang, "Secure switched observers for cyber-physical systems under sparse sensor attacks: A set cover approach," *IEEE Transactions on Automatic Control*, vol. 64, no. 9, pp. 3949–3955, 2019.

[12] X. Liu, Y. Mo, and E. Garone, "Local decomposition of kalman filters and its application for secure state estimation," *IEEE Transactions on Automatic Control*, pp. 1–1, 2020.

[13] Y. Mao, A. Mitra, S. Sundaram, and P. Tabuada, "On the computational complexity of the secure state-reconstruction problem," 2021. [Online]. Available: https://arxiv.org/abs/2101.01827

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

[15] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the royal statistical society series b-methodological*, vol. 58, pp. 267–288, 1996.