Attack-Resilient State Estimation for Noisy Dynamical Systems

Miroslav Pajic, Member, IEEE, Insup Lee, Fellow, IEEE, and George J. Pappas, Fellow, IEEE

Abstract—Several recent incidents have clearly illustrated the susceptibility of cyber-physical systems (CPS) to attacks, raising attention to security challenges in these systems. The tight interaction between information technology and the physical world has introduced new vulnerabilities that cannot be addressed with the use of standard cryptographic security techniques. Accordingly, the problem of state estimation in the presence of sensor and actuator attacks has attracted significant attention in the past. Unlike the existing work, in this paper we consider the problem of attack-resilient state estimation in the presence of bounded-size noise. We focus on the most general model for sensor attacks where any signal can be injected via compromised sensors. Specifically, we present an \( l_0 \)-based state estimator that can be formulated as a mixed-integer linear program and its convex relaxation based on the \( l_1 \) norm. For both attack-resilient state estimators, we derive rigorous analytic bounds on the state-estimation errors caused by the presence of noise. Our analysis shows that the worst-case error is linear with the size of the noise, and thus the attacker cannot exploit the noise to introduce unbounded state-estimation errors. Finally, we show how the \( l_0 \) and \( l_1 \)-based attack-resilient state estimators can be used for sound attack detection and identification; we provide conditions on the size of attack vectors that ensure correct identification of compromised sensors.

Index Terms—Attack-resilient state estimation, robustness of state-estimators, cyber-physical systems security, linear systems.

I. INTRODUCTION

M ost of existing control systems have not been built with security in mind. Even with the proliferation of different networking technologies and the use of more open control architectures, until recently security of control systems has usually been an afterthought. Yet, with the advance of cyber-physical systems (CPS), the tight interaction between information technology and the physical world have made control components of CPS vulnerable to attack vectors well beyond the standard cyber attacks [3]. In the last few years, several incidents have clearly illustrated susceptibility of CPS to attacks and raised attention to security challenges in these systems. These include attacks on large-scale systems such as the Maroochy Water breach [4] and the StuxNet virus attack on an industrial SCADA system [5], [6]. In addition, attacks on modern vehicles [7], [8], [9] and RQ-170 Sentinel US drone that was captured in Iran [10], [11], show that even widely used, safety-critical automotive and avionics CPS can be compromised by malicious attackers.

A typical cyber-physical system contains an internal network, or multiple networks, connected by a gateway to external communications network. In many cases, such as in automotive industry, systems rely on perimeter security where internal networks are resource constrained, mostly depending on the security of the gateway and external communication channels. However, the gateway may be compromised, becoming a threat to the system’s operation [7], [8]. In addition, some of the internal components might be tampered with, allowing the attacker to access to the internal communication network [9]. From the control of CPS perspective, attacks on the internal network, where the attacker inserts messages anywhere in the sensors-to-controllers-to-actuators pathway, can be modeled as additional signals injected into the control loop via system sensors and actuators [12]. While some of these attacks can be avoided with the use of standard cryptographic tools that guarantee integrity and authentication of data, this incurs a significant design, computational, and operational overhead for these usually resource-constrained systems.

On the other hand, unlike standard cyber systems, reported CPS vulnerabilities include non-invasive attacks on system sensors, where an adversarial signal is injected into the measured data by modifying a sensor’s physical environment. This has been illustrated with several non-invasive attacks on GPS-based navigation systems [13], [14], [15], and Anti-lock Braking Systems [16] in vehicles. These attacks show that the use of standard encryption and data-authentication based network security techniques does not guarantee secure control of CPS. In such cases, even if the stream of sensor data is properly encrypted, it would still contain wrong values. Consequently, there is a need to focus on attack-resilient control of CPS, to ensure safety in such scenarios.

A. Related Work

In recent years significant efforts have been invested into development of control techniques that exploit some knowledge
of system dynamics for attack detection and attack-resilient control (e.g., [17], [12], [18], [19], [20], [21], [22], [23]). One line of work has focused on attack-detection [24], [25], [22], [23] based on the use of standard residual probability based detectors (e.g., chi-square detector). For example, in [22] the authors illustrate how these detectors can be used to detect integrity attacks on SCADA systems, while in [23] the authors focus on the design of watermarked control inputs for active attack detection.

In addition to attack-detection, the problem of state estimation in the presence of sensor and actuator attacks has attracted significant attention due to the fact that CPS capable of correctly estimating the plant’s state from corrupted measurements would be able to continue operating even under attack. For noiseless linear time-invariant (LTI) systems for which the exact plant model is known, the attack-resilient state estimation problem has been formulated as an \( l_0 \) optimization problem [18], [19]. In addition, in [26], the authors present a SMT-based state estimation technique.

However, it is unclear how robust are these state-estimators to noise and modeling errors; specifically, what kind of guarantees can be provided for performance of attack-resilient state estimators for noisy dynamical systems. To the best of our knowledge, the first work on this topic was [2]. In that paper, we introduced an \( l_0 \)-based attack-resilient state estimator for systems with bounded noise that can be formulated as a Mixed-Integer Linear Program (MILP). We also showed its robustness to noise and modeling errors, and provided a complex design-time procedure to bound the worst-case state estimation error in the presence of attacks.

It is worth noting that our work exploits some of the ideas initially introduced in the domain of compressed sensing (e.g., see [27] and the references within), starting from the problem considered in [28], [29] where a sparse state was to be extracted for noisy non-dynamical systems with a predefined measurement matrix and without any structured interference. These works were extended to the problem of the extraction of block-sparse signals for these systems in the presence of noise (e.g., [30]). On the other hand, error bounds for the estimation of (non-block) sparse signals in the presence of structured interference for noisy non-dynamical systems have been recently addressed in [31], [32]. Specifically, [31] considers systems with Gaussian measurement matrices, while [32] provides a very conservative error bound due to the fact that the authors assume that both state and interference are sparse.

B. Contributions of this Work

In this paper, we focus on the problem of attack-resilient state estimation for linear dynamical systems with bounded-size noise. We consider the most general model for sensor attacks where any signals can be injected via the compromised sensors [12]. We start from the \( l_0 \)-based state estimation procedure introduced in [19], and show how it can be adapted for systems with noise. The main limitation of the \( l_0 \)-based state estimators is that solving them is NP-hard in general. Therefore, by exploiting properties of the \( l_1 \) norm we provide a computationally efficient, convex optimization based state estimation procedure for noisy dynamical systems.

Furthermore, unlike our work in [2] for the \( l_0 \) estimator, we derive rigorous analytic bounds on the state-estimation errors for both \( l_0 \) and \( l_1 \)-based state estimation procedures. We show that the worst-case error is linear with the size of the noise, and when the number of attacked sensors is not higher than a predefined number (that depends on the properties of the system’s observability matrix) the attacker cannot exploit noise and modeling errors to introduce unbounded state-estimation errors. In addition, we introduce a method that utilizes the presented attack-resilient state estimators for sound attack detection and identification, using the estimates of attack vectors provided by the estimators.

The rest of the paper is organized as follows: Section II introduces the problem formulation, while in Section III we present attack-resilient state estimation procedures based on the \( l_0 \) and \( l_1 \)-based state estimators, respectfully. Finally, in Section VI, we show how the presented state estimators can be used for sound attack detection and identification, followed by evaluation of the introduced robustness bounds (Section VII) and some concluding remarks in Section VIII.

C. Notation and Terminology

For a set \( S \), \( |S| \) denotes the cardinality (i.e., size) of the set. In addition, for a set \( K \subset S \), with \( K^D \) we denote the complement set of \( K \) with respect to \( S \) – i.e., \( K^D = S \setminus K \).

We use \( A^T \) to indicate the transpose of matrix \( A \), while \( i^{th} \) element of a vector \( x_k \) is denoted by \( x_{k,i} \). For vector \( x \) and matrix \( A \), we denote by \( |x| \) and \( |A| \) the vector and matrix whose elements are absolute values of the initial vector and matrix, respectively. Also, for matrices \( P \) and \( Q \), by \( P \leq Q \) we specify that the matrix \( P \) is element-wise smaller than the matrix \( Q \). In addition, for a symmetric matrix \( Q \), \( Q \preceq 0 \) denotes that the matrix is positive semidefinite.

We use \( \mathbb{R} \) to denote the set of reals. In addition, \( I_p \) denotes the identity matrix of size \( p \), while \( \mathbb{I}(\cdot) \) denotes the indicator function. Finally, for a vector \( e \in \mathbb{R}^p \), the support of the vector is the set

\[
\text{supp}(e) = \{i \mid e_i \neq 0\} \subseteq \{1, 2, ..., p\},
\]

while \( l_0 \) norm of vector \( e \) is the cardinality of \( \text{supp}(e) \) – i.e., \( \|e\|_{l_0} = |\text{supp}(e)| \).

II. PROBLEM DESCRIPTION

We consider LTI systems of the form:

\[
x_{k+1} = Ax_k \\
y_k = Cx_k + w_k + e_k.
\]

(1)

The plant’s output vector \( y \in \mathbb{R}^p \) contains measurements of the plant’s state \( x \in \mathbb{R}^n \) provided by \( p \) sensors from the set \( S = \{s_1, s_2, ..., s_p\} \). We assume the measurement noise vector \( w \in \mathbb{R}^p \) to be bounded; specifically, we assume that \( |w_k| \leq \delta_{w_k} \), for all \( k \geq 0 \). Finally, the sparse vector \( e \in \mathbb{R}^p \) with
support in the set $\mathcal{K} \subseteq \mathcal{S}$ denotes the attack vector injected by a malicious attacker using sensors from the set $\mathcal{K}$.

The attack-resilient state estimation problem focuses on reconstruction of the initial system state $x_0$ from a set of $N$ output observations $y_0, y_1, \ldots, y_{N-1}$. These observations are potentially corrupted by an attacker with access to the sensors from the set $\mathcal{K}$ i.e.,

$$y_k = CA^k x_0 + e_k + w_k.$$ 

One additional goal is to provide identification of the compromised sensors (i.e., identify sensors from $\mathcal{K}$), since the actual set $\mathcal{K}$ of compromised sensors is not known before the estimation.

1) Model Motivation: The aforementioned attack-resilient state estimation problem can be also used for the general form of LTI systems:

$$x_{k+1} = Ax_k + Bu_k + v^p_k$$

$$y_k = Cx_k + v^m_k + e_k.$$ 

Here, $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, and $C \in \mathbb{R}^{p \times n}$, while process and measurement noise, $v^p \in \mathbb{R}^p$ and $v^m \in \mathbb{R}^p$ respectively, are bounded in size. In this general case, to obtain the plant’s state at any time-step $t$ (i.e., $x_t$), the goal is to utilize the previous $N$ sensor measurement vectors $(y_{t-N+1}, \ldots, y_t)$ and actuator inputs $(u_{t-N+1}, \ldots, u_{t-1})$ to evaluate the state $x_{t-N+1}$.

For dynamical systems without noise, the state can be obtained as the minimization argument of the following optimization problem [19], [2]

$$\min_{E_t,N \in \mathbb{R}^{p \times N}} \|E_t,N\|_{l_0} \quad s.t. \quad E_t,N = Y_{t,N} - \Phi_N(x)$$

Here, the matrix $E_{t,N} = [e_{t-N+1} \mid e_{t-N+2} \mid \ldots \mid e_t]$ captures the last $N$ attacks vectors. In addition, $Y_{t,N} = [y_{t-N+1} \mid y_{t-N+2} \mid \ldots \mid y_t]$ maintains the last $N$ sensor measurements compensated for the impacts of the inputs applied during that interval i.e.,

$$\hat{y}_k = y_k, \quad k = t - N + 1$$

$$\hat{y}_k = y_k - \sum_{i=0}^{k-t+N-2} CA^iBu_{k-1-i}, \quad k = t - N + 2, \ldots, N.$$

Finally, linear mapping $\Phi_N : \mathbb{R}^n \rightarrow \mathbb{R}^{p \times N}$ defined as $\Phi_N(x) = [Cx \mid CAx \mid \ldots \mid CA^{N-1}x]$ specifies the observed system evolution, due to its dynamics, from initial state $x$.

Consequently, even for the general form of LTI systems as in (2), the problem of state-estimation can be mapped into the state estimation for systems from (1), where control inputs are discarded. Furthermore, as shown in [2], the bounds on the size of measurement noise in (2) can be related to the bounds on the size of process and measurement noise vectors, $v^p$ and $v^m$.

1In this work, we sometimes abuse notation by using $\mathcal{K}$ to denote both the set of compromised sensors and the set of indices of the compromised sensors.

2We consider the measurement history size $N$ as an input parameter to the state-estimation procedure.

III. ATTACK-RESILIENT STATE ESTIMATORS

We start by introducing the following notation. We use $P_\mathcal{K}$ to denote the projection from the set $\mathcal{S}$ to set $\mathcal{K}$ by keeping only rows of $C$ with indices that correspond to sensors from $\mathcal{K}$. Formally, $P_{\mathcal{K}} = [s_{k_1} \mid \ldots \mid s_{k_{|\mathcal{K}|}}] \subseteq \mathcal{S}$ and $k_1 < k_2 < \ldots < k_{|\mathcal{K}|}$, and $i^T_{\mathcal{K}}$ denotes the row vector (of the appropriate size) with a 1 in its $j^{th}$ position being the only nonzero element of the vector. Furthermore, for any sensor $s_i$ and set $\mathcal{K}$ we define the matrices $O_{s_i}$ and $O_{\mathcal{K}}$ as

$$O_{s_i} = \begin{bmatrix} P_{\{s_i\}}C \\ P_{\{s_i\}}CA^{N-1} \end{bmatrix}, \quad O_{\mathcal{K}} = \begin{bmatrix} O_{s_1} \\ \vdots \end{bmatrix}$$

We will also slightly abuse the notation by using $O_i$ to denote $O_{s_i}$ for each sensor $s_i$.

In addition, we use $e_i = [e_{s_1} \mid e_{s_2} \mid \ldots \mid e_{s_{N-1}}]^T \in \mathbb{R}^N$, for all $i \in \{1, \ldots, p\}$, to denote the values injected via sensor $s_i$ (i.e., attack signals on sensor $s_i$) at time-steps $0, \ldots, N - 1$. From the definition, if $s_i \notin \mathcal{K}$ then $e_i = 0 \in \mathbb{R}^N$. Similarly, for all $i \in \{1, \ldots, p\}$, we use $\tilde{y}_i = [y_{s_1} \mid y_{s_2} \mid \ldots \mid y_{s_{N-1}}]^T \in \mathbb{R}^N$ and $\tilde{w}_i = [w_{s_1} \mid w_{s_2} \mid \ldots \mid w_{s_{N-1}}]^T \in \mathbb{R}^N$ to denote all measurements obtained by the sensor $s_i$ and measurement noise at the sensor respectively, at time-steps $0, \ldots, N - 1$. Hence, we have that for all $1 \leq i \leq p$:

$$\tilde{y}_i = O_i x_0 + e_i + \tilde{w}_i$$

Finally, we define block vectors $\tilde{y}, \tilde{e}, \tilde{w} \in \mathbb{R}^{pN}$ as $\tilde{y} = [\tilde{y}_1 \mid \ldots \mid \tilde{y}_p]^T$, $\tilde{e} = [\tilde{e}_1 \mid \ldots \mid \tilde{e}_p]^T$, and $\tilde{w} = [\tilde{w}_1 \mid \ldots \mid \tilde{w}_p]^T$, and matrix $O = [O_1 \mid \ldots \mid O_p]^T$. Since each element of the measurement noise vectors $w_{0i}, \ldots, w_{N-1}$ is bounded (i.e., $|w_{ki}| \leq \delta_{wi}, 0 \leq k \leq N - 1, 1 \leq i \leq p$), we denote by $O_{\mathcal{W}} \subseteq \mathbb{R}^{pN}$ the feasible set of noise vectors $\tilde{w}$. In addition, for any set $R \subseteq \mathcal{S}$, we define $\tilde{w}_R$ to be the block vector obtained by concatenating $\tilde{w}_{s_i}$ for all $s_i \in R$ starting from the smallest $i$ to the largest, while the corresponding $\Omega_{\mathcal{R}} \subseteq \mathbb{R}^{pN}$ denotes the feasible set of vectors $\tilde{w}_R$. We similarly define the matrix $O_{\mathcal{R}}$ to be obtained by concatenating matrices $O_i$ for all $s_i \in R$.

Now, from (5), it follows that

$$\tilde{y} = O_i x_0 + \tilde{e} + \tilde{w}.$$ 

For block vectors obtained by concatenating $p$ vectors, such as $\tilde{e}$ and $\tilde{y}$, we also use the notation from [30]

$$\|\tilde{e}\|_{l_2,l_0} = \sum_{i=1}^{p} \mathbb{I}(\|\tilde{e}_i\|_{l_2} > 0)$$

$$\|\tilde{e}\|_{l_2,l_1} = \sum_{i=1}^{p} |\tilde{e}_i|_{l_2}$$

Note that for any block vector $\tilde{e}$ it holds that $\|\tilde{e}\|_{l_2,l_0} = \|\tilde{e}\|_{l_2,l_1}$ for any $t \geq 1$. This allows us to define block q-sparse

$^1$It is worth nothing that the vector $\tilde{e}_i$ corresponds to the $i^{th}$ row of the matrix $E$ from (3).

$^2$The matrix $O$ is obtained by reordering rows of the standard observability matrix for the system $(A, C)$, and thus, it has the same rank as the observability matrix.
vector $\hat{e}$ as a vector that satisfies $\|\hat{e}\|_{l_2,l_0} = q$, meaning that it has $q$ nonzero sub-vectors. Hence, if the set of compromised sensors $K$ has $q$ elements (i.e., $|K| = q$) then vector $\hat{e}$ is $q$-block sparse.

Using the above notation, the optimization problem (3) can be represented as:

$$P_0 : \quad \min_{\hat{e},x} \|\hat{e}\|_{l_2,l_0}$$

s. t. $\hat{y} - Ox_0 - \hat{e} = 0$ \hspace{1cm} (8)

Now, consider the measurement vector $\hat{y}$ for a noiseless system’s (i.e., when $\Omega = 0 \in \mathbb{R}^{pN}$) evolution due to the initial state $x_0$ and attack vector $\hat{e}^*$. If the number of attacked sensors $q = |K|$ is not higher than a certain number $q_{\text{max}}$,\footnote{The number $q_{\text{max}}$ depends on the properties of the observability matrix of the system. We will address this in more detail in Section IV.} the minimization arguments of the problem $P_0$ are exactly the initial state $x_0$ and the attack vector $\hat{e}^*$ [19]. Thus, in this case the estimator $P_0$ also correctly identifies the set of attacked sensors $K$. Furthermore, for noiseless systems $P_0$ is optimal in the sense that if another estimator can recover the initial state (which would also result in identification of the attacked sensors), the attack-resilient state estimator based on $P_0$ can as well [19].

On the other hand, $P_0$ cannot be used when noisy sensor measurements are available (i.e., when $\Omega \neq 0 \in \mathbb{R}^{pN}$). For instance, in this case the point $(x_0, \hat{e}^*)$ might not even be feasible. Thus, there is need to adapt problem $P_0$ to non-ideal models that capture system noise. To achieve this, we consider the following problem that relaxes the equality constraint from (8) by including a noise allowance

$$P_{0,\omega} : \quad \min_{\hat{e},x} \|\hat{e}\|_{l_2,l_0}$$

s. t. $\hat{y} - Ox_0 - \hat{e} = \hat{w}$

$\hat{w} \in \Omega$ \hspace{1cm} (9)

The problem $P_{0,\omega}$ involves combinatorial optimization that can be solved using MILP solvers. However, solving $P_{0,\omega}$ is NP-hard in the general case, which limits its use on smaller size systems. A common approach used in compressed sensing is to replace $l_0$ norm by $l_1$ norm, which effectively convexifies the problem and reduces its computational requirements. Consequently, to perform the attack-resilient state estimation we also consider the following optimization problem

$$P_{1,\omega} : \quad \min_{\hat{e},x} \|\hat{e}\|_{l_2,l_1}$$

s. t. $\hat{y} - Ox_0 - \hat{e} = \hat{w}$

$\hat{w} \in \Omega$ \hspace{1cm} (10)

However, it is unclear what guarantees can be provided regarding the performance of the attack-resilient state estimators $P_{0,\omega}$ and $P_{1,\omega}$. Specifically, we are interested in obtaining worst-case bounds on the state estimation errors caused by noise and attacks on sensors, and answering the question whether the attacker can exploit the noise to introduce an unbounded state estimation error. Furthermore, we will investigate conditions that ensure that the presented state estimators can be used to correctly identify the set of attacked sensors.

### IV. Performance Guarantees for $P_{0,\omega}$ Estimator

In this section, we focus on the performance degradation of the $P_{0,\omega}$ state estimator due to the existence of noise. Specifically, we are interested in providing bounds on $\Delta x^{l_0}$ that is defined as

$$\min_{q_0,\omega, \hat{e}^{l_0}} q_{0,\omega} = \|\hat{e}^{l_0}\|_{l_2,l_0}$$

$$\Delta x^{l_0} = x_{0,\omega} - x_0, \quad \Delta \hat{e}^{l_0} = \hat{e}^{l_0} - \hat{e}^*$$

We will also denote $\hat{e}^{l_0}$ and $\hat{e}^*$ as $\Delta x^{l_0}, \Delta \hat{e}^{l_0}$, and $\hat{e}^{l_0}$, respectively.

We consider systems where the number of compromised sensors $q = |K|$ is not higher than $q_{\text{max}}$ – the maximal number of attacked sensors for which the system’s state can be recovered in the noiseless case. Thus, before we proceed with our analysis, we first characterize conditions under which it is possible to perform the state estimation even for noiseless systems. We start with the following definition.

**Definition 1 ([33]):** An LTI system with the form as in (1) is said to be $s$-sparse observable if for every set $K \subseteq S$ of size $s$ (i.e., $|K| = s$), the pair $(A, P_{x0})$ is observable. \hspace{1cm} \Box

From the analysis in [19] the following holds.

**Lemma 1:** $q_{\text{max}}$ is equal to the maximal $s$ for which the system is $s$-sparse observable.

For considered systems, the following theorem provides a bound on the maximal state estimation error caused by the existence of noise.

**Theorem 1:** If $q$ sensors have been attacked, where $q \leq q_{\text{max}}$, then the error $\Delta x^{l_0}$ of the state estimate obtained from optimization problem $P_{0,\omega}$ satisfies

$$\|\Delta x^{l_0}\|_{l_2} \leq 2 \cdot \max_{R \subseteq S, |R| = p - 2q_{\text{max}}} \left(\|O_R^\parallel \|_{l_2} \cdot \max_{\hat{w} \in \Omega} \|\hat{w}_R\|_{l_2}\right)$$

where $O_R^\parallel$ denotes the pseudoinverse of $O_R$ (i.e., $O_R^\parallel = (O_R^\parallel)^{-1}O_R^\parallel$).

**Proof 1:** From (12) and the definition of $P_{0,\omega}$ it follows that $\|\Delta \hat{e}^{l_0} + \hat{e}^*\|_{l_2,l_0} \leq \|\hat{e}^*\|_{l_2,l_0}$. Since for all vectors $a, b,$ and $l_0 > 0$, $\|a + b\|_{l_2} \leq \|a\|_{l_2} + \|b\|_{l_2}$, we have that $\|\Delta \hat{e}^{l_0} + \hat{e}^*\|_{l_2,l_0} \leq \|\Delta \hat{e}^{l_0}\|_{l_2,l_0} + \|\hat{e}^*\|_{l_2,l_0}$. Therefore,

$$\|\Delta \hat{e}^{l_0}\|_{l_2,l_0} \leq 2\|\hat{e}^*\|_{l_2,l_0} r_1 \leq 2q_{\text{max}},$$

where $r_1$ holds because $\|\hat{e}^*\|_{l_2,l_0} = q$ and the number of attacked sensors $q$ is bounded by $q_{\text{max}}$.

From (6), we have that $\hat{e}^* = \hat{y} - Ox_{0,\omega} - \hat{w}^*$. Similarly, from the constraint (9) it follows that $\Delta \hat{e}^{l_0} = \hat{y} - Ox_{0,\omega} - \hat{w}^{l_0}$, which implies

$$\Delta \hat{e}^{l_0} = -O \Delta x^{l_0} - \hat{w}.$$
This implies that at least \( f = p - 2q_{\text{max}} \) blocks of \( \hat{z} \) are zero subvectors. Let’s denote their indexes as \( i_1, \ldots, i_f \), such that \( i_1 < \ldots < i_f \) and the set of sensors corresponding to these indexes as \( \mathcal{R} \) (i.e., \( \mathcal{R} = \{s_{i_1}, \ldots, s_{i_f}\} \)). Hence, we have that

\[
O_\mathcal{R} \Delta x^{l_0} = -\Delta \tilde{w}_\mathcal{R}
\]  

(16)

where \( \Delta \tilde{w}_\mathcal{R} = \tilde{w}_\mathcal{R}^{l_0} - \tilde{w}_\mathcal{R} \), with \( \tilde{w}_\mathcal{R}^{l_0}, \tilde{w}_\mathcal{R} \in \Omega_\mathcal{R} \).

Note that the set \( \mathcal{R} \) has \( f = p - 2q_{\text{max}} \) elements, and since the system is \( 2q_{\text{max}} \)-sparse observable (from Lemma 1), the pair \((A, P_{\mathcal{R}C})\) is observable (and \( f \geq 1 \)). Thus, the matrix \( O_\mathcal{R} \) is full (column) rank and we can define the pseudo-inverse matrix \( O_\mathcal{R}^+ = (O_\mathcal{R}^T O_\mathcal{R})^{-1} O_\mathcal{R}^T \), from which it follows that

\[
\Delta x^{l_0} = -O_\mathcal{R}^+ \Delta \tilde{w}_\mathcal{R} \Rightarrow \|\Delta x^{l_0}\|_2 \leq \|O_\mathcal{R}^+\|_2 \cdot \|\Delta \tilde{w}_\mathcal{R}\|_2 \Rightarrow \\
\|\Delta x^{l_0}\|_2 \leq \max_{\mathcal{R} \subseteq \mathcal{S}} \left( \|O_\mathcal{R}^+\|_2 \cdot \max_{\tilde{w}_\mathcal{R}^{l_0}, \tilde{w}_\mathcal{R} \in \Omega_\mathcal{R}} \|\tilde{w}_\mathcal{R}^+ - \tilde{w}_\mathcal{R}^{l_0}\|_2 \right)
\]

Since

\[
\max_{\tilde{w}_\mathcal{R}^{l_0}, \tilde{w}_\mathcal{R} \in \Omega_\mathcal{R}} \|\tilde{w}_\mathcal{R}^+ - \tilde{w}_\mathcal{R}^{l_0}\|_2 \leq 2 \max_{\tilde{w}_\mathcal{R} \in \Omega_\mathcal{R}} \|\tilde{w}_\mathcal{R}\|_2,
\]

we have that (13) is satisfied, which concludes the proof.

It is important to highlight that the bound on the right hand side of (13) is linear in the size of noise. Furthermore, the above theorem states that if at most \( q_{\text{max}} \) sensors have been compromised, the attacker cannot exploit the noise to introduce an unbounded state estimation error. Another thing to consider is the complexity of computing the term in (13). To determine the state estimation bound we need to check \((p - 2q_{\text{max}})\) different subsets \( \mathcal{R} \) of the set \( \mathcal{S} \), and for each \( \mathcal{R} \) compute

\[
\|O_\mathcal{R}^+\|_2 \cdot \max_{\tilde{w}_\mathcal{R} \in \Omega_\mathcal{R}} \|\tilde{w}_\mathcal{R}\|_2 = \lambda_{\text{max}}^{O_\mathcal{R}^+} \cdot \max_{\tilde{w}_\mathcal{R} \in \Omega_\mathcal{R}} \|\tilde{w}_\mathcal{R}\|_2,
\]

where \( \lambda_{\text{max}}^{O_\mathcal{R}^+} \) denotes the largest singular value of \( O_\mathcal{R}^+ \), and

\[
\max_{\tilde{w}_\mathcal{R} \in \Omega_\mathcal{R}} \|\tilde{w}_\mathcal{R}\|_2 = \sqrt{\sum_{k=0}^{N-1} \sum_{s_i \in \mathcal{R}} (\beta_{w,s_i})^2}
\]

for \( \Omega_\mathcal{R} \) defined as in Section III.\(^7\) This is significantly lower than the required computational cost for the robustness analysis from [2].

Finally, it is worth noting that for almost all systems (i.e., for almost all pairs of matrices \( A, C \)) we have that \( q_{\text{max}} = \lfloor p/2 - 1 \rfloor \) [19], meaning that \( 1 \leq p - 2q_{\text{max}} \leq 2 \). Thus, for almost all systems, to obtain the bound we would need to evaluate the above term for either \( p \) or \( p(p - 1)/2 \) sets \( \mathcal{R} \) only.

\(^7\)On the other hand, if the noise bounds in \( \Omega \) are defined as bounds on the \( l_2 \) norm of noise for each sensor at each time-step, this term would be equal to the sum of the squared norms.

V. ROBUSTNESS OF P\(_{1,\omega}\) ESTIMATOR TO NOISE

In this section, we provide a bound on the error of the \( P_{1,\omega} \) estimator due to noise. We start by introducing notation similar to the one used in the previous section:

\[
(x_{i,\omega}, \hat{e}_i) = \arg \min_{P_{1,\omega}} \Delta x_i = x_{i,\omega} - x_0, \quad \Delta \hat{e}_i = \hat{e}_i - e^*_i
\]

(17)

(18)

Specifically, we are interested in obtaining a bound on \( \Delta x_i \).

Theorem 2: When sensors from set \( \mathcal{K} \subset \mathcal{S} \) are attacked, state estimation error \( \hat{e}_i \) satisfies the following constraint

\[
\sum_{i \in \mathcal{K}} \|O_i \Delta x_i\|_2 \leq \sum_{i \in \mathcal{K}} \|O_i \Delta \hat{e}_i\|_2 + 2\sigma_{\mathcal{K}} \Omega,
\]

(19)

where \( \sigma_{\mathcal{K}} = \max_{\tilde{w}_\mathcal{K} \in \Omega_{\mathcal{K}}} \|\tilde{w}_\mathcal{K}\|_2 \).

Proof 2: Since \( \tilde{e}_i \) is a minimizer of the problem \( P_{1,\omega} \), it follows that \( \|\hat{e}_i\|_2 \leq \|\tilde{e}_i\|_2 \). Thus, to find a bound on the \( \Delta \hat{e}_i \) we consider a set that contains all feasible \( \Delta \hat{e}_i \), which is defined as:

\[
\{ \Delta \hat{e}_i \in \mathbb{R}^{N} \mid \|\Delta \hat{e}_i + \tilde{e}_i\|_2, \Delta \hat{e}_i \leq \|\tilde{e}_i\|_2, \tilde{e}_i \}
\]

(20)

By starting from the above feasibility condition, it follows that:

\[
\sum_{i=1}^{p} \|\Delta \hat{e}_i + \tilde{e}_i\|_2 \leq \sum_{i=1}^{p} \|\tilde{e}_i\|_2 \Rightarrow \\
0 \geq \sum_{i=1}^{p} \|\Delta \hat{e}_i + \tilde{e}_i\|_2 - \sum_{i=1}^{p} \|\tilde{e}_i\|_2 = \\
\sum_{s_i \in \mathcal{K}} \|\Delta \hat{e}_i + \tilde{e}_i\|_2 - \sum_{s_i \in \mathcal{K}} \|\tilde{e}_i\|_2 \geq \\
\sum_{s_i \in \mathcal{K}} \|\tilde{e}_i\|_2 - \sum_{s_i \in \mathcal{K}} \|\Delta \hat{e}_i\|_2 = \\
\sum_{s_i \in \mathcal{K}} \|\tilde{e}_i\|_2 - \sum_{s_i \in \mathcal{K}} \|\Delta \hat{e}_i\|_2 = \\
\sum_{i=1}^{p} \|\Delta \hat{e}_i\|_2 - 2 \sum_{s_i \in \mathcal{K}} \|\Delta \hat{e}_i\|_2.
\]

Here, \( r_1 \) is satisfied by the fact that only sensors from the set \( \mathcal{K} \) are attacked – i.e., all other blocks of the attack vector \( \tilde{e}_i \) are zero. Relation \( r_2 \) follows from the fact that \( ||a + b|| \geq ||a|| - ||b|| \), for any \( a, b \).\(^8\)

Thus, the set from (20) can be overapproximated by the set

\[
\left\{ \Delta \hat{e}_i \in \mathbb{R}^{N} \mid \|\Delta \hat{e}_i\|_2, \Delta \hat{e}_i \leq 2 \sum_{s_i \in \mathcal{K}} \|\Delta \hat{e}_i\|_2 \right\}
\]

(21)

From (5), we have that \( \tilde{e}_i = \tilde{y}_i - O_i x_0 - \tilde{w}_i \). Similarly, from (10) it follows that \( \tilde{e}^{l_1}_i = \tilde{y}_i - O_i x_{i,l_1} - \tilde{w}^{l_1}_i \), which implies

\[
\Delta \hat{e}^{l_1}_i = -(O_i \Delta x_{i,l_1} - \Delta \tilde{w}_i)
\]

(22)

where \( \Delta \tilde{w}_i = \tilde{w}_i - \bar{w}_i \) with \( \tilde{w}_i, \bar{w}_i \in \Omega_{\{s_i\}} \).

\(^8\)In this paper, when a norm is not clearly specified we imply that a statement is valid for any norm.
Hence, the constraint from the set definition in (21) can be represented as:

$$\sum_{i=1}^{p} \|\Delta \hat{e}_i\|_{l_2} \leq 2 \sum_{s_i \in K} \|\Delta \hat{e}_i\|_{l_2} \Leftrightarrow$$

$$\sum_{s_i \in K^C} \|O_i \Delta x_i^l + \Delta \hat{w}_i\|_{l_2} \leq \sum_{s_i \in K} \|O_i \Delta x_i^l + \Delta \hat{w}_i\|_{l_2}$$

(23)

Again, using properties of norms we have

$$\|O_i \Delta x_i^l + \Delta \hat{w}_i\|_{l_2} \geq \|O_i \Delta x_i^l\|_{l_2} - \|\Delta \hat{w}_i\|_{l_2}$$

$$\|O_i \Delta x_i^l + \Delta \hat{w}_i\|_{l_2} \leq \|O_i \Delta x_i^l\|_{l_2} + \|\Delta \hat{w}_i\|_{l_2},$$

and thus, from (23) it follows that the set constraint from (21) can be additionally relaxed to

$$\sum_{s_i \in K^C} \|O_i \Delta x_i^l\|_{l_2} \leq \sum_{s_i \in K} \|O_i \Delta x_i^l\|_{l_2} + \delta_w,$$

(24)

where

$$\delta_w = \sum_{i=0}^{p} \|\Delta \hat{w}_i\|_{l_2} \leq 2 \max_{i \in \Omega_{1}(s_j)} \sum_{i=0}^{p} \|\hat{w}_i\|_{l_2} = 2 \max_{i \in \Omega} \|\hat{w}_i\|_{l_2,l_1},$$

which concludes the proof.

Remark 1: Proposition 6 from [19] states that $P_{1,\omega}$ can correctly estimate the state for noiseless systems ($\Omega = 0$) if and only if for all $K$ such that $|K| = q$, it holds that:

$$\sum_{s_i \in K^C} \|O_i \Delta x_i\|_{l_2} > \sum_{s_i \in K} \|O_i \Delta x_i\|_{l_2}, \quad \forall x \in \mathbb{R}^n \setminus \{0\}.$$  

This means that (19) is tight for noiseless systems, since for $\Omega = 0$ (19) takes the form $\sum_{s_i \in K^C} \|O_i \Delta x_i\|_{l_2} \leq \sum_{s_i \in K} \|O_i \Delta x_i\|_{l_2}$; this constraint when combined with (25) implies that for noiseless systems $\Delta x_i = 0$, meaning that the state is correctly reconstructed.

Finally, if we consider systems that can deal with up to $q$ attacks when there is no noise, from (19) and (25) it follows that the feasible set for the state estimation vector $\Delta x_i$ can be described as the set where $\Delta x_i = 0$ or it satisfies

$$\sum_{s_i \in K^C} \|O_i \Delta x_i\|_{l_2} < \sum_{s_i \in K^C} \|O_i \Delta x_i\|_{l_2} \leq \sum_{s_i \in K} \|O_i \Delta x_i\|_{l_2} + 2\sigma_\Omega$$

for all $K \subset S$, such that $|K| = q$.

From the relationship between $l_2$ and $l_1$ norms where

$$\|\alpha\|_{l_2} \geq \|\alpha\|_{l_1} \geq \frac{1}{\sqrt{n}} \|\alpha\|_{l_1} \quad \forall \alpha \in \mathbb{R}^n,$$

(26)

it follows that $\|O_i \Delta x_i\|_{l_1} \geq \|O_i \Delta x_i\|_{l_2} \geq \frac{1}{\sqrt{n}} \|O_i \Delta x_i\|_{l_2}$. Therefore,

$$\sum_{s_i \in K^C} \|O_i \Delta x_i\|_{l_2} \geq \frac{1}{\sqrt{n}} \sum_{s_i \in K^C} \|O_i \Delta x_i\|_{l_1} = \frac{\|O_K \Delta x_i\|_{l_1}}{\sqrt{n}}$$

(27)

$$\sum_{s_i \in K^C} \|O_i \Delta x_i\|_{l_2} \leq \sum_{s_i \in K} \|O_i \Delta x_i\|_{l_1} = \|O_K \Delta x_i\|_{l_1}.$$  

The above inequalities along with Theorem 2 prove the following corollary.

**Corollary 1:** When sensors from set $K \subset S$ are attacked, the state estimation error $\Delta x_i$ satisfies

$$\|O_K \Delta x_i\|_{l_1} \leq \sqrt{n} \|O_K \Delta x_i\|_{l_1} + 2\sqrt{n} \sigma_\Omega.$$

where $\sigma_\Omega = \max_{i \in \Omega} \|\hat{w}_i\|_{l_2,l_1}.$

Both conditions from Theorem 2 and Corollary 1 define sets that contain all feasible $\Delta x_i$ when less than or equal to $q$ sensors are attacked; the case where $q_1 < q$ sensors are attacked is covered by the scenario where $|K| = q$ sensors are attacked, but $q - q_1$ sensors are inserting zero signals. However, maximization problems over these sets may be hard to solve in the general case. Thus, we introduce the following theorem that for a special class of systems provides an analytic formula for $\|\Delta x_i\|_{l_2}$.

**Theorem 3:** Suppose that for all $K \subset S$ with $|K| = q$ it holds

$$O_K^T e O_K q N^2 O_K^T O_K \geq \lambda I_n$$

(28)

for some $\lambda > 0$. Then if at most $q$ nodes are compromised the following condition holds:

$$\|\Delta x_i\|_{l_2} \leq \frac{2\sqrt{n} \sigma_\Omega}{\lambda} \max_{K \subset S, |K| = q} (\|O_K\|_{l_2} + \sqrt{q} N \|O_K\|_{l_2})$$

(29)

**Proof 3:** We start by assuming that the set of compromised sensors $K$ has $q$ elements. From (27) and (26) it follows that

$$\|O_K \Delta x_i\|_{l_2} \leq \sqrt{q} N \|O_K \Delta x_i\|_{l_2} + 2\sqrt{n} \sigma_\Omega.$$

(30)

Let’s denote $T = \|O_K \Delta x_i\|_{l_2} - \sqrt{q} N \|O_K \Delta x_i\|_{l_2}$. Then, it holds that for $\Delta x_i \neq 0$:

$$T = \frac{\|O_K \Delta x_i\|_{l_2}^2 - q N^2 \|O_K \Delta x_i\|_{l_2}^2}{\|O_K \Delta x_i\|_{l_2} + \sqrt{q} N \|O_K \Delta x_i\|_{l_2}} = \frac{\|O_K \Delta x_i\|_{l_2}^2 + \sqrt{q} N \|O_K \Delta x_i\|_{l_2}}{\lambda \|\Delta x_i\|_{l_2}}$$

(31)

Here, $r_1$ holds due to (28), while $r_2$ is caused by the fact that $\|Ax\|_{l_2} \leq \|A\|_{l_2} \|x\|_{l_2}$ for any matrix $A$ and vector $x$.

Therefore, from (30) and (31), $\Delta x_i$ satisfies

$$\frac{\lambda}{\|O_K\|_{l_2} + \sqrt{q} N \|O_K\|_{l_2}} \|\Delta x_i\|_{l_2} \leq 2\sqrt{n} \sigma_\Omega$$

(32)

which implies (29).

On the other hand, consider a set $K_1 \subset S$ with $|K_1| < q$.

If (28) is satisfied for all $K$ with exactly $q$ sensors, then it also holds for $K_1$.

This case is also covered by the scenario where $q$ sensors including the sensors from $K_1$ are compromised, while the injected attack signals on $q - |K_1|$ sensors that are not in $K_1$ are equal to zero.

The error signal in this case satisfies (30), although effectively only $|K_1| < q$ are used to inject attacks, which concludes the proof.

Although Theorem 3 provides an analytic bound for the worst-case state estimation error obtained by $P_{1,\omega}$ for a certain class of systems, it is worth noting that it could heavily overapproximate the error due to the gains caused by the
of compromised sensors – i.e., that no valid sound for which the attack vectors estimates can be used for bound (29) (where \(\mathbb{E}_1\)).

In this section, we consider conditions for which the attack vectors estimates can be used for sound identification of compromised sensors – i.e., that no valid sensor will be identified as under attack.

An obvious candidate for identification procedure would be to use the policy that classifies sensor \(s_i\) as attacked if and only if \(\mathbb{E}_1\). Note that, unless we can guarantee that the set of identified attacked sensors is a subset of the actual set of attacked sensors \(\mathcal{K}\) (which is not known in advance), we cannot guarantee soundness of this identification procedure.

On the other hand, we can use the state estimation guarantees presented in the previous two sections to provide a sound attack identification procedure. Consider the state \(\mathbb{E}_1\). If \(\mathbb{E}_1\) (i.e., sensor \(s_i\) is not attacked), then \(\mathbb{E}_1\). Consequently, if there is a bound on the values for \(\mathbb{E}_1\) (i.e., error of the attack vector estimation for sensor \(s_i\)), we can guarantee that all attack vector estimates \(\mathbb{E}_1\) that violate the bound effectively correspond to scenarios where sensor \(s_i\) is attacked.

To determine this bound, referred to as \(D_1\), we use that from (15): \(\Delta e_i = -O_i \Delta x_i - \Delta w_i\). Thus,

\[
\|\Delta e_i\|_2 \leq \|O_i\|_2 \|\Delta x_i\|_2 + \|\Delta w_i\|_2 \\
\leq \|O_i\|_2 \|\Delta x_i\|_2 + 2 \max_{w_i \in \Omega_i} \|\Delta w_i\|_2. \tag{34}
\]

Therefore, the bounds for \(\|\Delta x_i\|_2\), which we will refer to as \(D_1\), can be used to compute a bound for \(\|\Delta e_i\|_2\) as follows

\[
D_1 = \|O_i\|_2 D_1 + 2 \max_{w_i \in \Omega_i} \|\Delta w_i\|_2.
\]

9In this section, we will use \(I_1\) notation (instead of \(I_0\) or \(I_1\)) whenever we describe results that hold for both \(P_{0,\omega}\) and \(P_{1,\omega}\) obtained estimates.

10To the best of our knowledge, even for a simpler problem of estimation of sparse signals \(\alpha_0\) from noisy measurements \(z\) obtained using an overcomplete dictionary \(\Phi\) (i.e., \(z = \Phi\alpha_0 + v\)), the \(l_0\) based solution [28], [29] does not guarantee correct support recovery for \(\alpha_0\).

For instance, when \(P_{0,\omega}\) is used, the bound \(D_1\) on \(\|\Delta e_i\|_2\) is

\[
D_1 = 2\|O_i\|_2 \max_{w_i \in \Omega_i} \|\Delta w_i\|_2.
\]

Now we can define a \(P_{1,\omega}\)-based \((t = 0, 1)\) attack identification scheme as:

\[
\text{Attacked}^t(s_i) = \{\|e_i\|_2 > D_1\}, \quad i = 1, \ldots, p. \tag{36}
\]

The following theorem shows soundness of the proposed attack identification scheme.

\[\text{Theorem 4: If \text{Attacked}^t(s_i) = 1 then sensor } s_i \in \mathcal{K}. \]

\[\text{Furthermore, for all attack vectors } \hat{e}^t \text{ for which } \|\hat{e}^t\|_2 > 2D_1, \text{ the attack on sensor } s_i \text{ will be correctly detected (i.e., Attacked}^t(s_i) = 1). \]

\[\text{Proof 4: Suppose Attacked}^t(s_i) = 1, \text{ implying that } \|\hat{e}^t\|_2 > D_1^t. \]

Thus, \(\|\hat{e}^t\|_2 > 0\), which is the actual attack vector on \(s_i\) is non-zero, which means that sensor \(s_i \in \mathcal{K}\).

On the other hand, let’s assume that \(\|\hat{e}^t\|_2 > 2D_1^t\). This implies the following:

\[2D_1^t < \|\hat{e}^t - \Delta e_i\|_2 \leq \|\hat{e}^t\|_2 + \|\Delta e_i\|_2 \leq 2D_1^t + \|\hat{e}^t\|_2.
\]

Hence, \(\|\hat{e}^t\|_2 > D_1^t\), and Attacked\(^t\)(s_i) = 1.

\[\text{VII. Evaluation}\]

To evaluate conservativeness of the state-estimation error bounds presented in this work, we exploit the evaluation approach from [2]. We randomly generated 100 systems with \(n = 10\) states and \(p = 5\) sensors, and 100 systems with \(n = 20\) states and \(p = 11\) sensors. For each of the 200 systems, we evaluated the state-estimation error \(\Delta x_i\) in 1000 experiments for various attack and noise realizations, where the number of attacked sensors was less than or equal to 2 for the systems with \(p = 5\) sensors, and less than or equal to 5 for systems with \(p = 11\) sensors.

The focus of our evaluation was the comparison between the state-estimation bounds and the observed state estimation errors due to the presence of noise. In both simulations and calculations of the error bounds we considered the case when the window size \(N\) is equal to the number of system states (i.e., \(N = n\)). The results of our evaluation are presented in Fig. 1 and Fig. 2. Fig. 1(a) and Fig. 2(a) present histograms of \(\Delta x_i\) errors for all 1000 attack scenarios for two randomly selected system with \(n = 10\) and \(n = 20\). As can be seen, the bound from Theorem 1 is an order of magnitude larger than the average state-estimation error for each system.

Furthermore, we investigated the ratio between the worst-case observed state estimation error for all 1000 simulations of...
Fig. 1. Simulation results for 1000 runs of 100 randomly selected systems with $n = 10$ states and $p = 5$ sensors.

(a) Histogram for a system with the precomputed error bound equal to 28.6

(b) Histogram of the maximal relative state-estimation error for all 100 system
each system $S$ – i.e., $\max_{i=1:1000} \|\Delta x_{i0}^S\|_2$, and the system’s
error bounds $D_{x0}^S$ from Theorem 1

$$\text{Rel}_{\text{error}}_S = \frac{\max_{i=1:1000} \|\Delta x_{i0}^S\|_2}{D_{x0}^S}.$$  

Histograms of the relative errors for both types of systems are shown in Fig. 1(b) and Fig. 2(b). As can be observed, the maximal observed state estimation error reaches 16% of the computed bound for smaller systems ($n = 10$ states), while for larger systems (with $n = 20$ states) the maximal relative error reaches 1.5% of the computed bounds.

Conservativeness of the presented results was partially caused by the fact that we only simulated random initial points and random attack vectors, where sensor attacks are generated independently for each attacked sensor and noise profiles. As a result, the considered scenarios clearly do not capture worst-case attacks (i.e., attacks that could maximize the state estimation errors); for each system, to obtain scenarios that result in the worst-case estimation errors it is necessary to derive the corresponding attack vectors (and the initial states), which is beyond the scope of this paper.

This has been highlighted in the discrepancy of the relative estimation errors for systems with different size, as illustrated in the histograms in Fig. 1(b) and Fig. 2(b). While simulating different attack and noise realizations, we observed that the obtained maximal relative estimation error reduces with an increase in the system size $n$ and hence the window size (since we analyzed systems for $N = n$). This can be explained by the fact that with the increase of the window size $N$, the number of attack vectors is also increased, meaning that due to the random attack vector selections, the probabilities to incorporate a worst-case attack is significantly reduced. However, for small systems (e.g., $n = 1, 2$ states) we were able to generate initial states and attack vectors for which the obtained bounds were tight – i.e., the error $\|\Delta x_{0}^S\|_2$ is equal to the obtained bounds.

Finally, the obtained bounds from Theorem 1 are only slightly more conservative than the bounds obtained using the procedure we introduced in [2]; for instance, for the systems with $n = 10$ states, the maximal relative error was 20%. On the other hand, the complexity of the error bounding algorithm from [2] limits its use to systems with smaller number of states, while the bound from Theorem 1 can be computed for
systems with \( n = 20 \) states and \( p = 11 \) sensors in the order of seconds.

VIII. CONCLUSION

In this paper, we have considered the problem of state estimation when some of the sensors are attacked by a malicious attacker. Unlike existing work on this topic, we have investigated the case when there is bounded-size noise in the system’s dynamics. We have shown how to use two estimators that incorporate noise allowance in its constraints (i.e., \( P_{0,\omega} \) and \( P_{1,\omega} \)) and proved that the worst-case state estimation error is linear with the size of the noise present in the system. The provided bounds illustrate that \( \delta_0 \) based state estimation results in significantly more accurate state estimation. However, the penalty is paid in the complexity of the procedure - \( P_{0,\omega} \) can be solved as a mixed integer linear program, which are NP hard in general, while \( P_{1,\omega} \) can be efficiently solved using standard convex solvers and is more suited for embedded control applications.

Finally, we have derived attack identification procedures, based on these estimators, that exploit the fact that besides state estimates, estimations of attack vectors are also provided. We have shown that the proposed attack identification schemes are sound, and derived conditions on signals injected via an attacked sensor that would guarantee identification of the compromised sensor. An avenue for future work would be to determine conditions when the support of estimated attack vectors is a subset of the set of attacked vectors.

REFERENCES


