2015 IEEE 54th Annual Conference on Decision and Control (CDC)
December 15-18, 2015. Osaka, Japan

Attack-Resilient State Estimation in the Presence of Noise

Miroslav Pajic Paulo Tabuada Insup Lee George J. Pappas

Abstract— **We consider the problem of attack-resilient state estimation in the presence of noise. We focus on the most general model for sensor attacks where any signal can be injected via the compromised sensors. 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 are presented. For both l_0 and l_1-based state estimators, we derive rigorous analytic bounds on the state-estimation errors. We show that the worst-case error is linear with the size of the noise, meaning that the attacker cannot exploit noise and modeling errors to introduce unbounded state-estimation errors. Finally, we show how the presented attack-resilient state estimators can be used for sound attack detection and identification, and provide conditions on the size of attack vectors that will ensure correct identification of compromised sensors.**

I. Introduction

In recent years, several incidents have raised attention to security challenges in existing control systems and illustrated their susceptibility to attacks. Examples of these incidents include the Maroochy Water breach [1], the StuxNet virus attack on an industrial SCADA system [2], and attacks on modern automotive systems [3]. Some of the documented control system vulnerabilities were exposed by non-invasive attacks on system sensors, where an adversarial signal is injected into the measured data by modifying a sensor's physical environment. For instance, several attacks on GPS based navigation systems (e.g., [4]) and Anti-lock Braking Systems [5] have been reported, illustrating that the use of standard authentication based network security techniques does not guarantee security of control systems.

Consequently, 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., [6], [7], [8], [9], [10], [11]). One line of work has focused on attack-detection [12], [13]. Furthermore, state estimation in presence of sensor and actuator attacks has attracted significant attention due to the fact that systems 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 [8], [9]. In addition, in [14], the authors present an SMT-based state estimation technique.

However, for systems with noise, it is unclear what kind of guarantees can be given regarding the performance of attack-resilient state estimators. To the best of our knowledge, the first work on this topic was [15] where we introduced an l_0-based attack-resilient state estimator for systems with bounded noise, which 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 sensor attacks.

In this paper, we focus on the problem of attack-resilient state estimation for linear dynamical systems with noise. We consider the most general model for sensor attacks where *any* signals can be injected via the compromised sensors [7]. We start from the l_0-based state estimation procedure introduced in [9] and show how it can be adapted for systems with noise. The main limitation of the l_0-based state estimators is that solving the corresponding optimization problem is NP-hard in general. Thus, by exploiting properties of the l_1 norm we provide a computationally efficient, convex optimization based state estimation procedure for systems with noise. We also 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, which depends on the properties of the system's observability matrix, the attacker cannot exploit noise and modeling errors to introduce unbounded state-estimation errors. Finally, we present how these attack-resilient state estimators can be exploited for sound attack detection and identification.

Note that our work exploits some of the ideas initially introduced in the domain of compressed sensing [16]. In particular, the problem of extraction of block-sparse signals have been recently addressed in the community (e.g., [17]), while [18] provides guarantees for extraction of (non-block) sparse signals in presence of structured interference.

A. Notation and Terminology

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

This material is based on research sponsored by DARPA under agreement number FA8750-12-2-0247. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of DARPA or the U.S. Government. This work was also supported in part by NSF CNS-1505701 grant and a grant from Intel.

M. Pajic is with the Department of Electrical and Computer Engineering, Durham, NC, USA 27708. Email: miroslav.pajic@duke.edu. P. Tabuada is with the Department of Electrical Engineering, UCLA, Los Angeles, CA, USA 90095. Email: tabuada@ee.ucla.edu. G. J. Pappas and I. Lee are with the Departments of Electrical and Systems Engineering and Computer and Information Science, respectively, University of Pennsylvania, Philadelphia, PA, USA 19104. Email: {pappasg,lee}@seas.upenn.edu.

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

We use $\mathbb{R}$ to denote the set of reals. Also, $\mathbf{I}_p$ denotes the identity matrix of size p, while $\mathbb{I}(\cdot)$ denotes the indicator function. Finally, for a vector $\mathbf{e} \in \mathbb{R}^p$, *the support* of the vector is the set $\text{supp}(\mathbf{e}) = \{i \mid \mathbf{e}_i \neq 0\} \subseteq \{1, 2, ..., p\}$, while l_0 norm of vector $\mathbf{e}$ is the cardinality of $\text{supp}(\mathbf{e})$ – i.e., $\|\mathbf{e}\|_{l_0} = |\text{supp}(\mathbf{e})|$.

II. Problem Description

We consider LTI systems of the form

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{A}\mathbf{x}_k \\ \mathbf{y}_k &= \mathbf{C}\mathbf{x}_k + \mathbf{w}_k + \mathbf{e}_k. \end{aligned} \tag{1}$$

The plant's output vector $\mathbf{y} \in \mathbb{R}^p$ contains measurements of the plant's state $\mathbf{x} \in \mathbb{R}^n$ provided by p sensors from the set $\mathcal{S} = \{s_1, s_2, ..., s_p\}$. We assume the measurement noise vector $\mathbf{w} \in \mathbb{R}^p$ to be bounded; specifically, we assume that $|\mathbf{w}_k| \leq \delta_{w_k}$, for all $k \geq 0$. Finally, the sparse vector $\mathbf{e} \in \mathbb{R}^p$ with support in set $\mathcal{K} \subseteq \mathcal{S}$ denotes the attack vector injected by a malicious attacker using sensors from $\mathcal{K}$.[1]

The attack-resilient state estimation problem focuses on reconstruction of the initial system state $\mathbf{x}_0$ from a set of N output observations[2] $\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_{N-1}$ corrupted by an attacker with access to the sensors from the set $\mathcal{K}$ – i.e.,

$$\mathbf{y}_k = \mathbf{C}\mathbf{A}^k\mathbf{x}_0 + \mathbf{e}_k + \mathbf{w}_k.$$

Since set $\mathcal{K}$ is not known before the estimation, additional goal is to identify compromised sensors (i.e., identify set $\mathcal{K}$).

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

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{v}^{\mathbf{p}}{}_k \\ \mathbf{y}_k &= \mathbf{C}\mathbf{x}_k + \mathbf{v}^{\mathbf{m}}{}_k + \mathbf{e}_k, \end{aligned} \tag{2}$$

with $\mathbf{A} \in \mathbb{R}^{n \times n}, \mathbf{B} \in \mathbb{R}^{n \times m}$, and $\mathbf{C} \in \mathbb{R}^{p \times n}$, while process and measurement noise, $\mathbf{v}^{\mathbf{p}} \in \mathbb{R}^n$ and $\mathbf{v}^{\mathbf{m}} \in \mathbb{R}^p$ respectively, are bounded in size. Here, to obtain the plant's state at any time-step t (i.e., $\mathbf{x}_t$), the goal is to utilize the previous N sensor measurement vectors $(\mathbf{y}_{t-N+1}, ..., \mathbf{y}_t)$ and actuator inputs $(\mathbf{u}_{t-N+1}, ..., \mathbf{u}_{t-1})$ to evaluate the state $\mathbf{x}_{t-N+1}$.

For noiseless systems, the state can be obtained as the minimization argument of the following optimization problem [9], [15]

$$\begin{aligned} \min_{\mathbf{E}_{t,N} \in \mathbb{R}^{p \times N},\ \mathbf{x} \in \mathbb{R}^n} \quad & \|\mathbf{E}_{t,N}\|_{l_0} \\ s.\ t. \quad & \mathbf{E}_{t,N} = \mathbf{Y}_{t,N} - \Phi_N(\mathbf{x}) \end{aligned} \tag{3}$$

[1]In this work, we sometimes abuse notation with $\mathcal{K}$ denoting both the set of compromised sensors and the set of indices of the compromised sensors.

[2]Note that the measurement history size N is an input parameter to the state-estimation procedure.

Here, the matrix $\mathbf{E}_{t,N} = [\mathbf{e}_{t-N+1}|\mathbf{e}_{t-N+2}|\ldots|\mathbf{e}_t]$ captures the last N attacks vectors. In addition, $\mathbf{Y}_{t,N} = [\tilde{\mathbf{y}}_{t-N+1}|\tilde{\mathbf{y}}_{t-N+2}|\ldots|\tilde{\mathbf{y}}_t]$ maintains the last N sensor measurements compensated for the impact of the inputs applied during that interval – i.e.,

$$\begin{aligned} \tilde{\mathbf{y}}_k &= \mathbf{y}_k, && k = t - N + 1 \\ \tilde{\mathbf{y}}_k &= \mathbf{y}_k - \sum_{i=0}^{k-t+N-2} \mathbf{C}\mathbf{A}^i\mathbf{B}\mathbf{u}_{k-1-i}, && k = t - N + 2, ..., N. \end{aligned}$$

Finally, the linear mapping $\Phi_N : \mathbb{R}^n \to \mathbb{R}^{p \times N}$ defined as $\Phi_N(\mathbf{x}) = \left[\mathbf{C}\mathbf{x}|\mathbf{C}\mathbf{A}\mathbf{x}|\ldots|\mathbf{C}\mathbf{A}^{N-1}\mathbf{x}\right]$ specifies the observed system evolution, due to its dynamics, from initial state $\mathbf{x}$.

Therefore, for the general form of LTI systems (2), the state-estimation problem can be mapped into the state estimation for systems from (1), where control inputs are discarded. In addition, as shown in [15], 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, $\mathbf{v}^{\mathbf{p}}$ and $\mathbf{v}^{\mathbf{m}}$.

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 $\mathbf{C}$ with indices that correspond to sensors from $\mathcal{K}$. Formally, $P_{\mathcal{K}} = \left[\mathbf{i}_{k_1} \cdots \mathbf{i}_{k_{|\mathcal{K}|}}\right]^T$, where $\mathcal{K} = \{s_{k_1}, ..., s_{k_{|\mathcal{K}|}}\} \subseteq \mathcal{S}$ and $k_1 < ... < k_{|\mathcal{K}|}$, and $\mathbf{i}_j^T$ denotes the row vector (of appropriate size) with a 1 in its j^{th} position being the only non-zero element of the vector. Also, for any sensor s_i we define the matrices $\mathbf{O}_{s_i}$ and $\mathbf{O}_{\mathcal{K}}$

$$\mathbf{O}_{s_i} = \begin{bmatrix} P_{\{s_i\}}\mathbf{C} \\ P_{\{s_i\}}\mathbf{C}\mathbf{A} \\ \vdots \\ P_{\{s_i\}}\mathbf{C}\mathbf{A}^{N-1} \end{bmatrix} \qquad \mathbf{O}_{\mathcal{K}} = \begin{bmatrix} \mathbf{O}_{s_{i_1}} \\ \mathbf{O}_{s_{i_2}} \\ \vdots \\ \mathbf{O}_{s_{i_{|\mathcal{K}|}}} \end{bmatrix}, \tag{4}$$

We will also slightly abuse the notation by using $\mathbf{O}_i$ to denote $\mathbf{O}_{s_i}$ for each sensor s_i.

In addition, we use $\tilde{\mathbf{e}}_i = \left[\mathbf{e}_{0,i}\ \mathbf{e}_{1,i}\ \cdots\ \mathbf{e}_{N-1,i}\right]^T \in \mathbb{R}^N$, for all $i \in \{1, ..., p\}$, to denote the values injected via sensor s_i (i.e., attack signals on sensor s_i) at time-steps $0, ..., N-1$.[3] From the definition, if $s_i \notin \mathcal{K}$ then $\tilde{\mathbf{e}}_i = \mathbf{0} \in \mathbb{R}^N$. Similarly, for all $i \in \{1, ..., p\}$, we use $\tilde{\mathbf{y}}_i = \left[\mathbf{y}_{0,i}\ \cdots\ \mathbf{y}_{N-1,i}\right]^T$ and $\tilde{\mathbf{w}}_i = \left[\mathbf{w}_{0,i}\ \ldots\ \mathbf{w}_{N-1,i}\right]^T$ to denote all measurements obtained by the sensor s_i and measurement noise at the sensor respectively, at time-steps $0, ..., N-1$. Hence, we have that for all $1 \leq i \leq p$

$$\tilde{\mathbf{y}}_i = \mathbf{O}_i\mathbf{x}_0 + \tilde{\mathbf{e}}_i + \tilde{\mathbf{w}}_i \tag{5}$$

Finally, we define block vectors $\tilde{\mathbf{y}}, \tilde{\mathbf{e}}, \tilde{\mathbf{w}} \in \mathbb{R}^{pN}$ as $\tilde{\mathbf{y}} = \left[\tilde{\mathbf{y}}_1^T\ \ldots\ \tilde{\mathbf{y}}_p^T\right]^T$, $\tilde{\mathbf{e}} = \left[\tilde{\mathbf{e}}_1^T\ \ldots\ \tilde{\mathbf{e}}_p^T\right]^T$, and $\tilde{\mathbf{w}} = \left[\tilde{\mathbf{w}}_1^T\ \ldots\ \tilde{\mathbf{w}}_p^T\right]^T$, and matrix $\mathbf{O} = \left[\mathbf{O}_1^T\ \ldots\ \mathbf{O}_p^T\right]^T$.[4] Since each element of the measurement noise vectors $\mathbf{w}_0, ..., \mathbf{w}_{N-1}$ is bounded (i.e., $|\mathbf{w}_{k,i}| \leq \delta_{w_{k,i}}, 0 \leq k \leq N-1, 1 \leq i \leq p$), we denote by $\Omega \subset \mathbb{R}^{pN}$ the feasible set of noise vectors $\tilde{\mathbf{w}}$.

[3]Note that vector $\tilde{\mathbf{e}}_i$ corresponds to the i^{th} row of the matrix $\mathbf{E}$ from (3).

[4]Since matrix $\mathbf{O}$ is obtained by reordering rows of the standard observability matrix $\mathbf{O}_{\mathcal{S}}$ for the system $(\mathbf{A}, \mathbf{C})$, $\text{rank}(\mathbf{O}) = \text{rank}(\mathbf{O}_{\mathcal{S}})$.

In addition, for any set $\mathcal{R} \subset \mathcal{S}$, we define $\tilde{\mathbf{w}}_{\mathcal{R}}$ to be the block vector obtained by concatenating $\tilde{\mathbf{w}}_{s_i}$ for all $s_i \in \mathcal{R}$ starting from the smallest i to the largest, while the corresponding $\Omega_{\mathcal{R}} \subset \mathbb{R}^{|\mathcal{R}|N}$ denotes the feasible set of vectors $\tilde{\mathbf{w}}_{\mathcal{R}}$. We similarly define the matrix $\mathbf{O}_{\mathcal{R}}$ to be obtained by concatenating matrices $\mathbf{O}_i$ for all $s_i \in \mathcal{R}$.

Now, from (5), it follows that

$$\tilde{\mathbf{y}} = \mathbf{O}\mathbf{x}_0 + \tilde{\mathbf{e}} + \tilde{\mathbf{w}}. \tag{6}$$

For block vectors obtained by concatenating p vectors, such as $\tilde{\mathbf{e}}$ and $\tilde{\mathbf{y}}$, we also use the notation from [17]

$$\begin{aligned} \|\tilde{\mathbf{e}}\|_{l_2,l_0} &= \sum_{i=1}^{p} \mathbb{I}(\|\tilde{\mathbf{e}}_i\|_{l_2} > 0) \\ \|\tilde{\mathbf{e}}\|_{l_2,l_1} &= \sum_{i=1}^{p} \|\tilde{\mathbf{e}}_i\|_{l_2} \end{aligned} \tag{7}$$

This allows us to define block q-sparse vector $\tilde{\mathbf{e}}$ as a vector that satisfies $\|\tilde{\mathbf{e}}\|_{l_2,l_0} = q$, meaning that it has q nonzero sub-vectors. Hence, if the set of compromised sensors $\mathcal{K}$ has q elements (i.e., $|\mathcal{K}| = q$) then vector $\tilde{\mathbf{e}}$ is q-block sparse.

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

$$\begin{aligned} P_0 : \quad & \min_{\tilde{\mathbf{e}},\mathbf{x}} \|\tilde{\mathbf{e}}\|_{l_2,l_0} \\ s.\ t. \quad & \tilde{\mathbf{y}} - \mathbf{O}\mathbf{x}_0 - \tilde{\mathbf{e}} = \mathbf{0} \end{aligned} \tag{8}$$

Now, consider the measurement vector $\tilde{\mathbf{y}}$ for a noiseless system's (i.e., when $\Omega = \mathbf{0} \in \mathbb{R}^{pN}$) evolution due to the initial state $\mathbf{x}_0$ and attack vector $\tilde{\mathbf{e}}^*$. If the number of attacked sensors $q = |\mathcal{K}|$ is not higher than a certain number q_{max},[5] the minimization arguments of the problem P_0 are exactly the initial state $\mathbf{x}_0$ and the attack vector $\tilde{\mathbf{e}}^*$ [9]. Thus, in this case the estimator P_0 also correctly identifies the set of attacked sensors $\mathcal{K}$. 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 [9].

On the other hand, P_0 cannot be used when noisy sensor measurements are available (i.e., when $\Omega \neq \mathbf{0} \in \mathbb{R}^{pN}$). For instance, in this case the point $(\mathbf{x}_0, \tilde{\mathbf{e}}^*)$ might not even be feasible. Thus, as we showed in [15], attack-resilient state estimation can be performed by solving the following problem that allows for the noise allowance

$$\begin{aligned} P_{0,\omega} : \quad & \min_{\tilde{\mathbf{e}},\mathbf{x}} \|\tilde{\mathbf{e}}\|_{l_2,l_0} \\ s.\ t. \quad & \tilde{\mathbf{y}} - \mathbf{O}\mathbf{x}_0 - \tilde{\mathbf{e}} = \tilde{\mathbf{w}} \\ & \tilde{\mathbf{w}} \in \Omega \end{aligned} \tag{9}$$

The problem $P_{0,\omega}$ involves combinatorial optimization and as we presented in [15] it 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

$$\begin{aligned} P_{1,\omega} : \quad & \min_{\tilde{\mathbf{e}},\mathbf{x}} \|\tilde{\mathbf{e}}\|_{l_2,l_1} \\ s.\ t. \quad & \tilde{\mathbf{y}} - \mathbf{O}\mathbf{x}_0 - \tilde{\mathbf{e}} = \tilde{\mathbf{w}} \\ & \tilde{\mathbf{w}} \in \Omega \end{aligned} \tag{10}$$

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 $\Delta\mathbf{x}^{l_0}$ that is defined as

$$(\mathbf{x}_{l_0,\omega}, \tilde{\mathbf{e}}^{l_0}) = arg\min P_{0,\omega}, \quad q_{0,\omega} = \|\tilde{\mathbf{e}}^{l_0}\|_{l_2,l_0} \tag{11}$$

$$\Delta\mathbf{x}^{l_0} = \mathbf{x}_{l_0,\omega} - \mathbf{x}_0, \qquad \Delta\tilde{\mathbf{e}}^{l_0} = \tilde{\mathbf{e}}^{l_0} - \tilde{\mathbf{e}}^* \tag{12}$$

We will also denote i^{th} blocks of $\Delta\mathbf{x}^{l_0}, \Delta\tilde{\mathbf{e}}^{l_0}$, and $\tilde{\mathbf{e}}^{l_0}$ as $\Delta\mathbf{x}_i^{l_0}, \Delta\tilde{\mathbf{e}}_i^{l_0}$, and $\tilde{\mathbf{e}}_i^{l_0}$, respectively.

We consider systems where the number of compromised sensors $q = |\mathcal{K}|$ is not higher than q_{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 ([19]): An LTI system from (1) is said to be s-sparse observable if for every set $\mathcal{K} \subset \mathcal{S}$ of size s (i.e., $|\mathcal{K}| = s$), the pair $(\mathbf{A}, P_{\mathcal{K}^0}\mathbf{C})$ is observable. □

From the analysis in [9] the following holds.

Lemma 1: q_{max} is equal to the maximal s for which the system is $2s$-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 $q \leq q_{max}$, then the error $\Delta\mathbf{x}^{l_0}$ of the state estimate obtained from optimization problem $P_{0,\omega}$ satisfies

$$\|\Delta\mathbf{x}^{l_0}\|_{l_2} \leq 2 \cdot \max_{\substack{\mathcal{R}\subset S,\\ |\mathcal{R}|=p-2q_{max}}} \left(\|\mathbf{O}_{\mathcal{R}}^{\dagger}\|_{l_2} \cdot \max_{\tilde{\mathbf{w}}_{\mathcal{R}} \in \Omega_{\mathcal{R}}} \|\tilde{\mathbf{w}}_{\mathcal{R}}\|_{l_2} \right) \tag{13}$$

where $\mathbf{O}_{\mathcal{R}}^{\dagger}$ denotes the pseudoinverse of $\mathbf{O}_{\mathcal{R}}$ (i.e., $\mathbf{O}_{\mathcal{R}}^{\dagger} = (\mathbf{O}_{\mathcal{R}}^T\mathbf{O}_{\mathcal{R}})^{-1}\mathbf{O}_{\mathcal{R}}^T$). □

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

[5] The number q_{max} depends on the properties of the observability matrix of the system. We will cover this in more detail in Section IV.

$\mathbb{I}(\|\mathbf{a}+\mathbf{b}\|_{l_2}>0) \geq \mathbb{I}(\|\mathbf{a}\|_{l_2}>0) - \mathbb{I}(\|\mathbf{b}\|_{l_2}>0)$,[6] we have that $\|\Delta\tilde{\mathbf{e}}^{l_0}+\tilde{\mathbf{e}}^*\|_{l_2,l_0} \geq \|\Delta\tilde{\mathbf{e}}^{l_0}\|_{l_2,l_0} - \|\tilde{\mathbf{e}}^*\|_{l_2,l_0}$. Therefore,

$$\|\Delta\tilde{\mathbf{e}}^{l_0}\|_{l_2,l_0} \leq 2\|\tilde{\mathbf{e}}^*\|_{l_2,l_0} \overset{r_1}{\leq} 2q_{max}, \tag{14}$$

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

From (6), we have that $\tilde{\mathbf{e}}^* = \tilde{\mathbf{y}} - \mathbf{O}\mathbf{x}_0 - \tilde{\mathbf{w}}^*$. Similarly, from the constraint (9) it follows that $\tilde{\mathbf{e}}^{l_0} = \tilde{\mathbf{y}} - \mathbf{O}\mathbf{x}_{l_0,\omega} - \tilde{\mathbf{w}}^{l_0}$, which implies

$$\Delta\tilde{\mathbf{e}}^{l_0} = -\mathbf{O}\Delta\mathbf{x}^{l_0} - \Delta\tilde{\mathbf{w}}. \tag{15}$$

Here, $\Delta\tilde{\mathbf{w}} = \tilde{\mathbf{w}}^{l_0} - \tilde{\mathbf{w}}^*$, with $\tilde{\mathbf{w}}^{l_0}, \tilde{\mathbf{w}}^* \in \Omega$.

Therefore, from (14) and (15), there exists an at most $2q_{max}$-sparse block vector $\tilde{\mathbf{z}} \in \mathbb{R}^{pN}$ – defined as $\tilde{\mathbf{z}} = -\Delta\tilde{\mathbf{e}}^{l_0}$, with at most $2q_{max}$ nonzero N-size blocks – such that

$$\mathbf{O}\Delta\mathbf{x}^{l_0} = -\Delta\tilde{\mathbf{w}} + \tilde{\mathbf{z}}.$$

This implies that at least $f = p - 2q_{max}$ blocks of $\tilde{\mathbf{z}}$ are zero subvectors. Let's denote their indexes as $i_1, ...i_f$, such that $i_1 < ... < i_f$ and the set of sensors corresponding to these indexes as $\mathcal{R}$ (i.e., $\mathcal{R} = \{s_{i_1}, ..., s_{i_f}\}$). Hence, we have that

$$\mathbf{O}_{\mathcal{R}}\Delta\mathbf{x}^{l_0} = -\Delta\tilde{\mathbf{w}}_{\mathcal{R}} \tag{16}$$

where $\Delta\tilde{\mathbf{w}}_{\mathcal{R}} = \tilde{\mathbf{w}}_{\mathcal{R}}^{l_0} - \tilde{\mathbf{w}}_{\mathcal{R}}^*$, with $\tilde{\mathbf{w}}_{\mathcal{R}}^{l_0}, \tilde{\mathbf{w}}_{\mathcal{R}}^* \in \Omega_{\mathcal{R}}$.

Set $\mathcal{R}$ has $f = p - 2q_{max}$ elements, and since the system is $2q_{max}$-sparse observable (from Lemma 1), it follows that the pair $(\mathbf{A}, P_{\mathcal{R}}\mathbf{C})$ is observable (and $f \geq 1$). Thus, the matrix $\mathbf{O}_{\mathcal{R}}$ is full (column) rank and we can define the pseudoinverse matrix $\mathbf{O}_{\mathcal{R}}^{\dagger} = (\mathbf{O}_{\mathcal{R}}^T\mathbf{O}_{\mathcal{R}})^{-1}\mathbf{O}_{\mathcal{R}}^T$, from which it follows that

$$\Delta\mathbf{x}^{l_0} = -\mathbf{O}_{\mathcal{R}}^{\dagger}\Delta\tilde{\mathbf{w}}_{\mathcal{R}} \Rightarrow \|\Delta\mathbf{x}^{l_0}\|_{l_2} \leq \|\mathbf{O}_{\mathcal{R}}^{\dagger}\|_{l_2} \cdot \|\Delta\tilde{\mathbf{w}}_{\mathcal{R}}\|_{l_2} \Rightarrow$$

$$\|\Delta\mathbf{x}^{l_0}\|_{l_2} \leq \max_{\substack{\mathcal{R}\subset S, |\mathcal{R}|=p-2q_{max} \\ \tilde{\mathbf{w}}_{\mathcal{R}}^{l_0}, \tilde{\mathbf{w}}_{\mathcal{R}}^* \in \Omega_{\mathcal{R}}}} \left(\|\mathbf{O}_{\mathcal{R}}^{\dagger}\|_{l_2} \cdot \|\tilde{\mathbf{w}}_{\mathcal{R}}^* - \tilde{\mathbf{w}}_{\mathcal{R}}^{l_0}\|_{l_2}\right)$$

$$\leq \max_{\substack{\mathcal{R}\subset S, \\ |\mathcal{R}|=p-2q_{max}}} \left(\|\mathbf{O}_{\mathcal{R}}^{\dagger}\|_{l_2} \cdot \max_{\tilde{\mathbf{w}}_{\mathcal{R}}^{l_0}, \tilde{\mathbf{w}}_{\mathcal{R}}^* \in \Omega_{\mathcal{R}}} \|\tilde{\mathbf{w}}_{\mathcal{R}}^* - \tilde{\mathbf{w}}_{\mathcal{R}}^{l_0}\|_{l_2}\right)$$

Since

$$\max_{\tilde{\mathbf{w}}_{\mathcal{R}}^{l_0}, \tilde{\mathbf{w}}_{\mathcal{R}}^* \in \Omega_{\mathcal{R}}} \|\tilde{\mathbf{w}}_{\mathcal{R}}^* - \tilde{\mathbf{w}}_{\mathcal{R}}^{l_0}\|_{l_2} \leq 2 \max_{\tilde{\mathbf{w}}_{\mathcal{R}} \in \Omega_{\mathcal{R}}} \|\tilde{\mathbf{w}}_{\mathcal{R}}\|_{l_2},$$

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. In addition, Theorem 1 states that if at most q_{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 $\binom{p}{p-2q_{max}}$ different subsets $\mathcal{R}$ of the set $\mathcal{S}$, and for each $\mathcal{R}$ compute

$$\|\mathbf{O}_{\mathcal{R}}^{\dagger}\|_{l_2} \cdot \max_{\tilde{\mathbf{w}} \in \Omega_{\mathcal{R}}} \|\tilde{\mathbf{w}}\|_{l_2} = \lambda_{max}^{\mathbf{O}_{\mathcal{R}}^{\dagger}} \cdot \max_{\tilde{\mathbf{w}}_{\mathcal{R}} \in \Omega_{\mathcal{R}}} \|\tilde{\mathbf{w}}_{\mathcal{R}}\|_{l_2},$$

where $\lambda_{max}^{\mathbf{O}_{\mathcal{R}}^{\dagger}}$ denotes the largest singular value of $\mathbf{O}_{\mathcal{R}}^{\dagger}$, and

$$\max_{\tilde{\mathbf{w}}_{\mathcal{R}} \in \Omega_{\mathcal{R}}} \|\tilde{\mathbf{w}}_{\mathcal{R}}\|_{l_2} = \sqrt{\sum_{s_i \in \mathcal{R}} \sum_{k=0}^{N-1} (\delta_{w_{k,i}})^2}$$

for $\Omega_{\mathcal{R}}$ defined as in Section III.[7] This is significantly lower than the required computational cost for the robustness analysis from [15].

Finally, for almost all systems (i.e., for almost all pairs of matrices $\mathbf{A}, \mathbf{C}$) we have that $q_{max} = \lceil p/2 - 1 \rceil$ [9], meaning that $1 \leq p - 2q_{max} \leq 2$. 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.

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:

$$(\mathbf{x}_{l_1,\omega}, \tilde{\mathbf{e}}^{l_1}) = arg\min P_{1,\omega} \tag{17}$$

$$\Delta\mathbf{x}^{l_1} = \mathbf{x}_{l_1,\omega} - \mathbf{x}_0, \quad \Delta\tilde{\mathbf{e}}^{l_1} = \tilde{\mathbf{e}}^{l_1} - \tilde{\mathbf{e}}^* \tag{18}$$

Specifically, we are interested in obtaining a bound on $\Delta\mathbf{x}^{l_1}$.

Theorem 2: When sensors from set $\mathcal{K} \subset S$ are attacked, state estimation error $\Delta\mathbf{x}^{l_1}$ satisfies the following constraint

$$\sum_{s_i \in \mathcal{K}^{\complement}} \|\mathbf{O}_i\Delta\mathbf{x}^{l_1}\|_{l_2} \leq \sum_{s_i \in \mathcal{K}} \|\mathbf{O}_i\Delta\mathbf{x}^{l_1}\|_{l_2} + 2\sigma_{\Omega}, \tag{19}$$

where $\sigma_{\Omega} = \max_{\tilde{\mathbf{w}} \in \Omega} \|\tilde{\mathbf{w}}\|_{l_2,l_1}$. □

Proof: The proof, which has been omitted due to space limitations, can be found in [20]. ■

Remark 1: Proposition 6 from [9] states that $P_{1,\omega}$ can correctly estimate the state for noiseless systems ($\Omega = \mathbf{0}$) if and only if for all $\mathcal{K}$ such that $|\mathcal{K}| = q$, it holds that:

$$\sum_{s_i \in \mathcal{K}^{\complement}} \|\mathbf{O}_i\mathbf{x}\|_{l_2} > \sum_{s_i \in \mathcal{K}} \|\mathbf{O}_i\mathbf{x}\|_{l_2}, \quad \forall\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}. \tag{20}$$

This implies that (19) is tight for noiseless systems, since for $\Omega = \mathbf{0}$, (19) takes the form $\sum_{s_i \in \mathcal{K}^{\complement}} \|\mathbf{O}_i\Delta\mathbf{x}^{l_1}\|_{l_2} \leq \sum_{s_i \in \mathcal{K}} \|\mathbf{O}_i\Delta\mathbf{x}^{l_1}\|_{l_2}$; this constraint when combined with (20) implies that for noiseless systems $\Delta\mathbf{x}^{l_1} = \mathbf{0}$ 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 (20) it follows that the feasible set for the state estimation vector $\Delta\mathbf{x}^{l_1}$ can be described as the set where $\Delta\mathbf{x}^{l_1} = \mathbf{0}$ or it satisfies

$$\sum_{s_i \in \mathcal{K}} \|\mathbf{O}_i\Delta\mathbf{x}^{l_1}\|_{l_2} < \sum_{s_i \in \mathcal{K}^{\complement}} \|\mathbf{O}_i\Delta\mathbf{x}^{l_1}\|_{l_2} \leq \sum_{s_i \in \mathcal{K}} \|\mathbf{O}_i\Delta\mathbf{x}^{l_1}\|_{l_2} + 2\sigma_{\Omega}$$

for all $\mathcal{K} \subset \mathcal{S}$, such that $|\mathcal{K}| = q$. □

From the relationship between l_2 and l_1 norms where

$$\|\alpha\|_{l_1} \geq \|\alpha\|_{l_2} \geq \frac{1}{\sqrt{n}}\|\alpha\|_{l_1}, \forall\alpha \in \mathbb{R}^n, \tag{21}$$

[6] Note that although l_0 is not convex, it satisfies the triangular inequality.

[7] On the other hand, if the noise bounds in Ω 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.

it follows that $\|\mathbf{O}_i\Delta\mathbf{x}^{l_1}\|_{l_1} \geq \|\mathbf{O}_i\Delta\mathbf{x}^{l_1}\|_{l_2} \geq \frac{1}{\sqrt{N}}\|\mathbf{O}_i\Delta\mathbf{x}^{l_1}\|_{l_1}$. Therefore,

$$\sum_{s_i\in\mathcal{K}^\complement}\|\mathbf{O}_i\Delta\mathbf{x}^{l_1}\|_{l_2} \geq \frac{1}{\sqrt{N}}\sum_{s_i\in\mathcal{K}^\complement}\|\mathbf{O}_i\Delta\mathbf{x}^{l_1}\|_{l_1} = \frac{\|\mathbf{O}_{\mathcal{K}^\complement}\Delta\mathbf{x}^{l_1}\|_{l_1}}{\sqrt{N}}$$

$$\sum_{s_i\in\mathcal{K}}\|\mathbf{O}_i\Delta\mathbf{x}^{l_1}\|_{l_2} \leq \sum_{s_i\in\mathcal{K}}\|\mathbf{O}_i\Delta\mathbf{x}^{l_1}\|_{l_1} = \|\mathbf{O}_{\mathcal{K}}\Delta\mathbf{x}^{l_1}\|_{l_1}.$$

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

Corollary 1: When sensors from set $\mathcal{K}\subset S$ are attacked, the state estimation error $\Delta\mathbf{x}^{l_1}$ satisfies

$$\|\mathbf{O}_{\mathcal{K}^\complement}\Delta\mathbf{x}^{l_1}\|_{l_1} \leq \sqrt{N}\|\mathbf{O}_{\mathcal{K}}\Delta\mathbf{x}^{l_1}\|_{l_1} + 2\sqrt{N}\sigma_\Omega. \quad (22)$$

where $\sigma_\Omega = \max_{\tilde{\mathbf{w}}\in\Omega}\|\tilde{\mathbf{w}}\|_{l_2,l_1}$. □

Both conditions from Theorem 2 and Corollary 1 define sets that contain all feasible $\Delta\mathbf{x}^{l_1}$ when less than or equal to q sensors are attacked.[8] 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 $\|\Delta\mathbf{x}^{l_1}\|_{l_2}$.

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

$$\mathbf{O}_{\mathcal{K}^\complement}^T\mathbf{O}_{\mathcal{K}^\complement} - qN^2\mathbf{O}_{\mathcal{K}}^T\mathbf{O}_{\mathcal{K}} \succeq \lambda\mathbf{I}_n \quad (23)$$

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

$$\|\Delta\mathbf{x}^{l_1}\|_{l_2} \leq \frac{2\sqrt{N}\sigma_\Omega}{\lambda}\cdot\max_{\mathcal{K}\subset\mathcal{S},|\mathcal{K}|=q}(\|\mathbf{O}_{\mathcal{K}^\complement}\|_{l_2} + \sqrt{q}N\|\mathbf{O}_{\mathcal{K}}\|_{l_2})$$

□

Proof: The proof, which has been omitted due to space limitations, can be found in [20]. ■

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 could heavily overapproximate the error due to the gains caused by the conversions between the norms (i.e., factor $\sqrt{q}N$). Still, along with Theorem 2 and Corollary 1, it provides the first analytic relation showing that the worst-case error is linear with the size of the noise, as in the case for the $P_{0,\omega}$ estimator.

VI. Attack Identification in Presence of Noise

In addition to computing a state estimate, the presented attack-resilient state estimation procedures also estimate attack vectors injected at time steps $k = 0, 1, ..., N-1$ (i.e., vectors $\tilde{\mathbf{e}}^{l_t}, t = 0, 1$[9]). Therefore, 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 $\mathbb{I}(\tilde{\mathbf{e}}_i^{l_t} \neq \mathbf{0})$. Note that, unless we can guarantee that the set of identified attacked sensors is a subset of the actual set of attacked sensors $\mathcal{K}$, we cannot guarantee soundness of this identification procedure.[10] 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 vector $\Delta\tilde{\mathbf{e}}^{l_t}$, denoting the errors of the obtained attack vector estimations for all sensors. If $\tilde{\mathbf{e}}_i^* = \mathbf{0}$ (i.e. sensor s_i is not attacked), then $\Delta\tilde{\mathbf{e}}_i^{l_t} = \tilde{\mathbf{e}}_i^{l_t}$. Hence, if there is a bound on the values for $\Delta\tilde{\mathbf{e}}_i^{l_t}$, we can guarantee that all attack vector estimates $\tilde{\mathbf{e}}_i^{l_t}$ that violate the bound effectively correspond to scenarios where sensor s_i is attacked.

To determine this bound, referred to as $D_i^{\tilde{\mathbf{e}}^{l_t}}$, we use that from (15) it follows that $\Delta\tilde{\mathbf{e}}_i^{l_t} = -\mathbf{O}_i\Delta\mathbf{x}^{l_t} - \Delta\tilde{\mathbf{w}}_i$. Thus,

$$\begin{aligned}\|\Delta\tilde{\mathbf{e}}_i^{l_t}\|_{l_2} &\leq \|\mathbf{O}_i\|_{l_2}\|\Delta\mathbf{x}^{l_t}\|_{l_2} + \|\Delta\tilde{\mathbf{w}}_i\|_{l_2}\\ &\leq \|\mathbf{O}_i\|_{l_2}\|\Delta\mathbf{x}^{l_t}\|_{l_2} + 2\max_{\tilde{\mathbf{w}}_i\in\Omega_{\{s_i\}}}\|\tilde{\mathbf{w}}_i\|_{l_2}.\end{aligned} \quad (24)$$

Thus, the bounds for $\|\Delta\mathbf{x}^{l_t}\|_{l_2}$, which we will refer to as $D^{\mathbf{x}^{l_t}}$, can be used to compute a bound for $\|\Delta\tilde{\mathbf{e}}_i^{l_t}\|_{l_2}$ as

$$D_i^{\tilde{\mathbf{e}}^{l_t}} = \|\mathbf{O}_i\|_{l_2}D^{\mathbf{x}^{l_t}} + 2\max_{\tilde{\mathbf{w}}_i\in\Omega_{\{s_i\}}}\|\tilde{\mathbf{w}}_i\|_{l_2}.$$

For example, when $P_{0,\omega}$ is used, the bound $D_i^{\tilde{\mathbf{e}}^{l_0}}$ on $\|\Delta\tilde{\mathbf{e}}_i^{l_t}\|_{l_2}$ is derived using $D^{\mathbf{x}^{l_0}}$ from Theorem 1 (i.e., Eq. (13)).

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

$$Attacked^{l_t}(s_i) = \mathbb{I}(\|\tilde{\mathbf{e}}_i^{l_t}\|_{l_2} > D_i^{\tilde{\mathbf{e}}^{l_t}}), \quad i = 1, ..., p. \quad (25)$$

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

Theorem 4: If $Attacked^{l_t}(s_i) = 1$ then sensor $s_i \in \mathcal{K}$. Furthermore, for all attack vectors $\tilde{\mathbf{e}}^*$ for which $\|\tilde{\mathbf{e}}_i^*\|_{l_2} > 2D_i^{\tilde{\mathbf{e}}^{l_t}}$, the attack on sensor s_i will be correctly detected (i.e., $Attacked^{l_t}(s_i) = 1$). □

Proof: Suppose $Attacked^{l_t}(s_i) = 1$, implying that $\|\tilde{\mathbf{e}}_i^{l_t}\|_{l_2} > D_i^{\tilde{\mathbf{e}}^{l_t}}$. Then,

$$D_i^{\tilde{\mathbf{e}}^{l_t}} < \|\Delta\tilde{\mathbf{e}}_i^{l_t} + \tilde{\mathbf{e}}_i^*\|_{l_2} \leq \|\Delta\tilde{\mathbf{e}}_i^{l_t}\|_{l_2} + \|\tilde{\mathbf{e}}_i^*\|_{l_2} \leq D_i^{\tilde{\mathbf{e}}^{l_t}} + \|\tilde{\mathbf{e}}_i^*\|_{l_2}.$$

Thus $\|\tilde{\mathbf{e}}_i^*\|_{l_2} > 0$, meaning that the actual attack vector on s_i is non-zero and sensor $s_1 \in \mathcal{K}$.

On the other hand, let's assume that $\|\tilde{\mathbf{e}}_i^*\|_{l_2} > 2D_i^{\tilde{\mathbf{e}}^{l_t}}$. This implies the following:

$$2D_i^{\tilde{\mathbf{e}}^{l_t}} < \|\tilde{\mathbf{e}}_i^{l_t} - \Delta\tilde{\mathbf{e}}_i^{l_t}\|_{l_2} \leq \|\tilde{\mathbf{e}}_i^{l_t}\|_{l_2} + \|\Delta\tilde{\mathbf{e}}_i^{l_t}\|_{l_2} \leq \|\tilde{\mathbf{e}}_i^{l_t}\|_{l_2} + D_i^{\tilde{\mathbf{e}}^{l_t}}.$$

Hence, $\|\tilde{\mathbf{e}}_i^{l_t}\|_{l_2} > D_i^{\tilde{\mathbf{e}}^{l_t}}$, and $Attacked^{l_t}(s_i) = 1$. ■

VII. Evaluation

Due to space limit, in this section we only discuss conservativeness of the derived l_0-based state estimation bound by considering 100 randomly generated dynamical systems with $n = 10$ states and $p = 5$ sensors, as it was done in [15]. More thorough evaluation of the attack-resilient state estimation

[8] Note that the case where $q_1 < q$ sensors are attacked is also covered by the scenario where $|\mathcal{K}|$ sensors are compromised, but $q - q_1$ sensors are inserting zero signals. Thus, it is enough to check the sets for $|\mathcal{K}| = q$ only.

[9] In this section, we will use l_t notation (instead of l_0 or l_1) whenever we describe results that hold for both $P_{0,\omega}$ and $P_{1,\omega}$ obtained estimates.

[10] To the best of our knowledge, even for a simpler problem of estimation of sparse signals α_0 from noisy measurements $\mathbf{z}$ obtained using an overcomplete dictionary $\mathbf{\Phi}$ (i.e., $\mathbf{z} = \Phi\alpha_0 + \mathbf{v}$), the l_0-based solution [16], [21] does not guarantee correct support recovery for α_0.

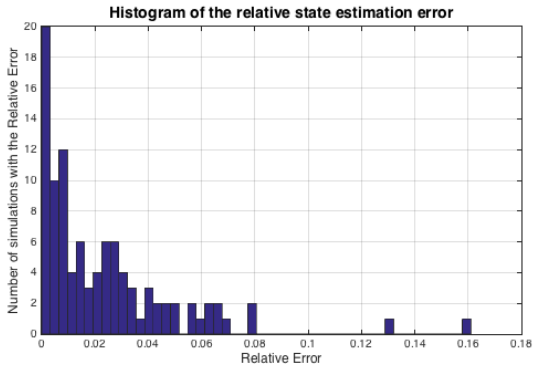

Fig. 1. Histogram of the maximal relative state-estimation error obtained from 1000 runs of 100 randomly selected systems with $n = 10$ states and $p = 5$ sensors.

bounds can be found in [20]. For each of the 100 systems, we evaluated the state-estimation error $\Delta \mathbf{x}^{l_0}$ in 1000 simulations for various attack and noise realizations, where the number of attacked sensors was less than or equal to 2. Finally, we considered the case where the window size $N = n$. Our main focus during the evaluation was the ratio between the worst-case observed state estimation error for all 1000 simulations of each system $\mathfrak{S}$ – i.e., $\max_{i=1:1000} \|\Delta \mathbf{x}^{l_0}_{\mathfrak{S}}\|_2$, and the system's error bounds $D^{\mathbf{x}^{l_0}}_{\mathfrak{S}}$ from Theorem 1.

A histogram of the relative errors for these systems is shown in Fig. 1, and as can be seen, the maximal observed state estimation error reaches 16% of the computed bound. Conservativeness of the presented results is partially caused by the fact that we only simulated random initial points and random attack vectors, which does not result in the worst-case estimation errors. 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.

VIII. Conclusion

In this paper, we have considered the problem of state estimation when some of system sensors are compromised by a malicious attacker. Unlike existing work on this topic, we have investigated the case when there is 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 l_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 optimization solvers and is more suited for embedded control applications.

Finally, we have derived attack identification procedures, based on these estimators. 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

[1] J. Slay and M. Miller, "Lessons learned from the maroochy water breach," in *Critical Infrast. Protection*, 2007, pp. 73–82.

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

[3] S. Checkoway, D. McCoy, B. Kantor, D. Anderson, H. Shacham, S. Savage, K. Koscher, A. Czeskis, F. Roesner, and T. Kohno, "Comprehensive experimental analyses of automotive attack surfaces," in *USENIX Security*, 2011.

[4] D. Shepard, J. Bhatti, and T. Humphreys, "Drone hack: Spoofing attack demonstration on a civilian unmanned aerial vehicle," *GPS World*, 2012.

[5] Y. Shoukry, P. Martin, P. Tabuada, and M. Srivastava, "Non-invasive spoofing attacks for anti-lock braking systems," in *CHES: Crypt. Hard. and Emb. Syst.*, 2013, pp. 55–72.

[6] R. Smith, "A decoupled feedback structure for covertly appropriating networked control systems," *Proc. IFAC World Congress*, pp. 90–95, 2011.

[7] A. Teixeira, D. Pérez, H. Sandberg, and K. H. Johansson, "Attack models and scenarios for networked control systems," in *Proc. the 1st international conference on High Confidence Networked Systems*, ser. HiCoNS '12, 2012, pp. 55–64.

[8] F. Pasqualetti, F. Dorfler, and F. Bullo, "Attack detection and identification in cyber-physical systems," *IEEE Trans. Autom. 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 Trans. Autom. Control*, vol. 59, pp. 1454–1467, 2014.

[10] S. Sundaram, M. Pajic, C. Hadjicostis, R. Mangharam, and G. Pappas, "The Wireless Control Network: Monitoring for malicious behavior," in *49th IEEE Conference on Decision and Control (CDC)*, 2010, pp. 5979–5984.

[11] F. Miao, M. Pajic, and G. Pappas, "Stochastic game approach for replay attack detection," in *52nd IEEE Annual Conference on Decision and Control (CDC)*, 2013, pp. 1854–1859.

[12] Y. Mo, T.-H. Kim, K. Brancik, D. Dickinson, H. Lee, A. Perrig, and B. Sinopoli, "Cyber–physical security of a smart grid infrastructure," *Proc. IEEE*, vol. 100, pp. 195–209, 2012.

[13] C. Kwon, W. Liu, and I. Hwang, "Security analysis for cyber-physical systems against stealthy deception attacks," in *American Control Conference (ACC)*, 2013, pp. 3344–3349.

[14] Y. Shoukry, A. Puggelli, P. Nuzzo, A. Sangiovanni-Vincentelli, S. A. Seshia, and P. Tabuada, "Sound and complete state estimation for linear dynamical systems under sensor attacks using satisfiability modulo theory solving," to appear.

[15] M. Pajic, J. Weimer, N. Bezzo, P. Tabuada, O. Sokolsky, I. Lee, and G. Pappas, "Robustness of attack-resilient state estimators," in *ACM/IEEE International Conference on Cyber-Physical Systems (ICCPS)*, 2014, pp. 163–174.

[16] D. L. Donoho, M. Elad, and V. N. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Trans. Inf. Theory*, vol. 52, pp. 6–18, 2006.

[17] Y. C. Eldar, P. Kuppinger, and H. Bolcskei, "Block-sparse signals: Uncertainty relations and efficient recovery," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3042–3054, 2010.

[18] R. Foygel and L. Mackey, "Corrupted sensing: Novel guarantees for separating structured signals," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 1223–1247, 2014.

[19] Y. Shoukry and P. Tabuada, "Event-triggered state observers for sparse sensor noise/attacks," *arXiv:1309.3511*, 2013.

[20] M. Pajic, P. Tabuada, I. Lee, and G. Pappas, "Attack-resilient state estimation for noisy dynamical systems," Tech. Rep., '15.

[21] J. A. Tropp, "Just relax: Convex programming methods for identifying sparse signals in noise," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 1030–1051, 2006.