Convex Optimization Based State Estimation against Sparse Integrity Attacks

Duo Han*, Yilin Mo* and Lihua Xie*

Abstract

We consider the problem of robust estimation in the presence of integrity attacks. There are m sensors monitoring the state and p of them are under attack. The malicious measurements collected by the compromised sensors can be manipulated arbitrarily by the attacker. The classical estimators such as the least squares estimator may not provide a reliable estimate under the so-called (p, m)-sparse attack. In this work, we are not restricting our efforts in studying whether any specific estimator is resilient to the attack or not, but instead we aim to present some generic sufficient and necessary conditions for robustness by considering a general class of convex optimization based estimators. The sufficient and necessary conditions are shown to be tight, with a trivial gap. We further specialize our result to scalar sensor measurements case.

I. Introduction

The concept of networks has been increasingly prevailing for decades, e.g., computer networks, sensor networks or social networks. Regardless of numerous benefits introduced by bridging machines or humans through networks, the interconnect and distributed nature renders networks vulnerable to various kinds of attacks, ranging from physical attacks to internet viruses to groundless rumors through online social networks. This article is concerned with the integrity attacks in sensor networks which are widely embedded in various industrial systems such as smart grid [1] or Supervisory Control And Data Acquisition (SCADA) systems [2]. During the integrity attack, the adversary can take full control of a subset of sensors and arbitrarily manipulate their measurements. The motivations for launching such an attack in industrial systems may include

$*$: School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. Email: dhanaa,ylmo,elhxie@ntu.edu.sg

creating arbitrage opportunities in electricity market, stealing gas or oil without being noticed, posing potential threat to national defense, etc. Since the first SCADA system malware (called Stuxnet) was discovered and extensively investigated [3], [4], increasing research attention has been paid to resolve the security issues in estimation and control systems [5].

In this article, we focus on the problem of robust estimation against compromised sensory data in order to mitigate the damage caused by the integrity attack. Robustness for an estimator is urgently needed since quite a number of the commonly used estimators under attack fail to give a reliable estimate and thus lead to poor system performance. For instance, a linear estimator is not robust since one bad measurement is enough to ruin the final estimate. A better estimator may be the geometric median of all measurements [6]. To be concrete, we consider the problem of estimating a vector state $x \in \mathbb{R}^n$ from measurements collected by m sensors, where the measurements are subject to any random noise. For practical reasons, the spatially distributed sensors cannot be fully guaranteed to be secure. Some of them may be controlled by the attacker and due to the resource limitation the attacker can only attack up to $p < m$ sensors. Without posing any restrictions on the attacker, we assume that the compromised sensory data can be arbitrarily changed.

Related Work: A quite similar problem in the context of power systems is bad data detection, which has been studied over the past decades [7], [8]. The method of checking the magnitude of residue is useful for identifying random bad data or outliers but may not work for intentional integrity attacks [9], [10]. For example, Liu et al. [11] successfully showed that a stealthy attack changing the state while not being detected is possible. Kim et al. [12] studied a so-called framing attack. Under such a attack, the bad data detector is misled to delete those critical measurements, without which the network is unobservable and a convert attack may be launched.

For dynamical systems, detecting malicious components via fault detection and isolation based methods has also been extensively studied, [13]–[17]. However, in most of these works, the system is assumed to be noiseless, which greatly favors the failure detector. Pajic et al. [18] improved the work by considering the systems with bounded noise. On the top of sufficient conditions for exact recovery in noiseless case, they showed that the worst error is still bounded even under attack. However, their estimator is based on a combinatorial optimization problem, which in general is computational hard to solve and may not be applicable for large scale systems. In [19], [20], the authors use reachability analysis and ellipsoid approximation to characterize all possible biases the adversary can inject to the system.

In the area of statistics, the concept of robust estimators is not new [21]–[23]. The robustness is often measured by breakdown points [24], [25] or influence functions [26]. Many existing works studied one or several estimators and discussed the breakdown point properties [27]–[30]. However, a unified analysis for most useful estimators is still absent.

Motivated by different behaviors of various estimators under the integrity attacks, we manage to provide a unified robustness analysis framework integrating most commonly used estimators. To reach this goal, we first give a formal definition on the robustness of an estimator. To achieve greater generality, a general convex optimization based estimator is proposed and necessary and sufficient conditions on the robustness of such an estimator is proved. The significance of this work is that the analytical results presented in this manuscript can be used for characterizing and designing a robust estimator in the presence of compromised sensory data.

The rest of the paper is organized as follows. In Section II we formulate the robust estimation problem. Our main results on the robustness of a general convex optimization based estimator is presented in Section III. We specialize our results for scalar sensor case in Section IV. The concluding remarks are given in Section V.

II. Problem Setup

A. System Model

Assume that m sensors are measuring the state x and the measurement equation for the ith sensor is given by

$$z_i = H_i x + w_i, \tag{1}$$

where $x \in \mathbb{R}^n$ is the state of interest, $z_i \in \mathbb{R}^{m_i}$ is the "true" measurement collected by the ith sensor, and $w_i \in \mathbb{R}^{m_i}$ is the measurement noise for the ith sensor. The measurement matrix $H \triangleq [H_1^\top, H_2^\top, \ldots, H_i^\top]^\top \in \mathbb{R}^{(\sum_i m_i)\times n}$ is assumed to be observable, *i.e.*, H is full column rank. In the presence of attacks, the measurement equation can be written as

$$y_i = z_i + a_i = H_i x + w_i + a_i, \tag{2}$$

where $y_i \in \mathbb{R}^{m_i}$ is the "manipulated" measurement and $a_i \in \mathbb{R}^{m_i}$ is the attack vector. In other words, the attacker can change the measurement of the ith sensor by a_i. Denote

$$\begin{aligned} z &\triangleq [z_1^\top, z_2^\top, \ldots, z_m^\top]^\top, & y &\triangleq [y_1^\top, y_2^\top, \ldots, y_m^\top]^\top, \\ w &\triangleq [w_1^\top, w_2^\top, \ldots, w_m^\top]^\top, & a &\triangleq [a_1^\top, a_2^\top, \ldots, a_m^\top]^\top. \end{aligned} \tag{3}$$

Denote the index set of all sensors as $\mathcal{S} \triangleq \{1, 2, \ldots, m\}$. For any index set $\mathcal{I} \subseteq \mathcal{S}$, define the complement set to be $\mathcal{I}^c \triangleq \mathcal{S} \backslash \mathcal{I}$. In our attack model, we assume that the attacker can only compromise at most p sensors but can arbitrarily choose a_i. Formally, a (p, m)-sparse attack can be defined as

Definition 1 ((p, m)-sparse attack): A vector a is called a (p, m)-sparse attack if there exists an index set $\mathcal{I} \subset \mathcal{S}$, such that:

(i) $\|a_i\| = 0, \ \forall i \in \mathcal{I}^c$;

(ii) $|\mathcal{I}| \leq p$.

Define the collection of a possible index set of malicious sensors as

$$\mathbb{C} \triangleq \{\mathcal{I} : \mathcal{I} \subset \mathcal{S}, |\mathcal{I}| = p\}.$$

The set of all possible (p, m)-sparse attacks is denoted as

$$\mathcal{A} \triangleq \bigcup_{\mathcal{I} \in \mathbb{C}} \{a : \|a_i\| = 0, i \in \mathcal{I}^c\}.$$

The main task of this work is to investigate the generic sufficient and necessary conditions for an estimator to be robust to (p, m)-sparse attacks. To this end, we first formally define the robustness of an estimator.

Definition 2 (Robustness): An estimator $g : \mathbb{R}^{(\sum_i m_i) \times n} \mapsto \mathbb{R}^n$ which maps the measurements y to a state estimate $\hat{x}$ is said to be robust to the (p, m)-sparse attack if it satisfies the following condition:

$$\|g(z) - g(z + a)\| \leq \mu(z), \ \forall a \in \mathcal{A}, \tag{4}$$

where $\mu : \mathbb{R}^{(\sum_i m_i) \times n} \mapsto \mathbb{R}$ is a real-valued mapping on z.

The robustness implies that the disturbance on the state estimate caused by an arbitrary attack is bounded. A trivial robust estimator is $g(y) = 0$ which provides very poor estimate. Therefore, another desirable property for an estimator is translation invariance, which is defined as follows:

Definition 3 (Translation invariance): An estimator g is translation invariant if $g(z + Hu) = u + g(z),\ \forall u \in \mathbb{R}^n$.

Remark 1: Notice that if an estimator is robust and translation invariant, then

$$\begin{aligned}\|g(z) - g(z+a)\| &= \|x + g(w) - x + g(w+a)\| \\ &= \|g(w) - g(w+a)\| \le \mu(w).\end{aligned}$$

Therefore, the maximum bias that can be injected by an adversary is only a function of the noise w.

In the next subsection, we propose a general convex optimization based estimator which is translation invariant.

B. A General Estimator

A large variety of estimators are developed by the research community to solve the state estimation problem. In order to achieve greater generality, we first propose a general convex optimization based estimator. We then show that many estimators can be rewritten in this general framework.

The estimator that we study in this paper is assumed to have the following form:

$$\hat{x} = g(y) \triangleq \arg\min_{\hat{x}} \sum_{i \in \mathcal{S}} f_i(y_i - H_i \hat{x}), \tag{5}$$

where the following properties of function $f_i : \mathbb{R}^{m_i} \mapsto \mathbb{R}$ are assumed:

(i) f_i is convex.
(ii) f_i is symmetric, *i.e.*, $f_i(u) = f_i(-u)$.
(iii) f_i is non-negative and $f_i(0) = 0$.

Remark 2: It is easy to check that the estimator g is translation invariant. One can view $y_i - H_i\hat{x}$ as the residue for the ith sensor and f_i as a cost function. The convex constraints on f_i ensures that the minimization problem can be solved in an efficient (possibly also distributed) way. The symmetric assumption on f_i is typically true for many practically used estimator and can actually be relaxed. The last assumption implies that the cost achieves minimum value when the residue is 0.

We now investigate several commonly used estimator and show that they can be written as (5).

(a) Least Square Estimator:

$$\hat{x} = \arg\min_{\hat{x}} \|y - H\hat{x}\|_2^2 = \arg\min_{\hat{x}} \sum_{i \in \mathcal{S}} \|y_i - H_i\hat{x}\|_2^2$$
$$= (H^\top H)^{-1} H^\top y. \tag{6}$$

(b) Another example is an estimator which minimizes the sum of the l_1 norm of the residue, *i.e.*,

$$\hat{x} = \arg\min_{\hat{x}} \sum_{i \in \mathcal{S}} \|y_i - H_i\hat{x}\|_1 . \tag{7}$$

In the case that $m_i = n$ and $H_i = \mathrm{I}_n,\ \forall i$, the estimate is a vector in which the ith entry is the median over the ith entries of all measurements y_i's.

(c) The following is designed to minimize the sum of the l_2 norm of the residue:

$$\hat{x} = \arg\min_{\hat{x}} \sum_{i \in \mathcal{S}} \|y_i - H_i\hat{x}\|_2 . \tag{8}$$

The optimal estimate in the case that $m_i = n$ and $H_i = \mathrm{I}_n,\ \forall i$ is the geometric median of all y_i's, which is called an L_1 estimator in [6]. In other words, $\hat{x}$ is the point in $\mathbb{R}^n$ that minimizes the sum of Euclidean distances from y_i to that point.

(d) Pajic et al. [18] proposed the following robust estimator in the presence of integrity attack:

$$\begin{aligned} &\underset{\hat{x},a,w}{\textit{minimize}} && \|w\|^2 \\ &\text{subject to} && y = H\hat{x} + w + a,\ \|a\|_0 \le q. \end{aligned}$$

However, the minimization problem involves zero-norm, and thus is difficult to solve in general. A commonly adopted approach is to use L_1 relaxation to approximate zero-norm, which leads to the following minimization problem:

$$\begin{aligned} &\underset{\hat{x},a,w}{\textit{minimize}} && \|w\|^2 + \lambda\|a\|_1 \\ &\text{subject to} && y = H\hat{x} + w + a. \end{aligned} \tag{9}$$

If we define the following function:

$$d(u) \triangleq \underset{a_i}{\textit{minimize}} \qquad \|u - a_i\|_2^2 + \lambda\, \|a_i\|_1 \tag{10}$$

Then one can easily prove that the optimization problem (9) can be rewritten as

$$\hat{x} = \arg\min_{\hat{x}} \sum_{i \in \mathcal{I}} d(y_i - H_i \hat{x}). \tag{11}$$

Apparently, the linear estimator (6) cannot give an estimate with bounded error even when only one measurement is arbitrarily manipulated. For other estimators, their robustness has been proved for some special cases. In the next section, we shall present sufficient and necessary conditions for the robustness of the general estimator (5). Since (7), (8) and (11) are all special cases of (5), we can easily analyze their individual robustness.

III. Robust Analysis for a General Estimator

This section is devoted to the derivation of necessary and sufficient conditions for the robustness of the general estimator. Denote the compact set $\mathcal{U} \triangleq \{u \in \mathbb{R}^n : \|u\| = 1\}$. Before proceeding to the main results, we need the following lemma.

Lemma 1: Let $q : \mathbb{R} \to \mathbb{R}$ be a convex function and $q(0) = 0$, then $q(t)/t$ is monotonically non-decreasing on $t \in \mathbb{R}^+$. Moreover,

$$q(t+1) - q(t) \geq q(t)/t. \tag{12}$$

For any $0 < \alpha < 1$, we have

$$q(\alpha t) \leq \alpha q(t) + q(0) = \alpha q(t).$$

Divide both side by αt, we can prove that $q(t)/t$ is monotonically non-decreasing. Therefore, $q(t+1)/(t+1) \geq q(t)/t$, which implies (12). As a consequence of Lemma 1, we know that $f_i(tH_iu)/t$ is monotonically non-decreasing. As a result, there are only two possibilities:

(i) $f_i(tH_iu)/t$ is bounded for all i and for all u, which implies that the limit $\lim_{t\to\infty} f_i(tH_iu)/t$ exists.

(ii) $f_i(tH_iu)/t$ is unbounded for some i and u.

The next lemma provides several important properties for the case where $\lim_{t\to\infty} f_i(tH_iu)/t$ exists, whose proof is reported in the appendix:

Lemma 2: If the following limit is well defined, *i.e.*, finite, for all $u \in \mathbb{R}^n$:

$$\lim_{t\to\infty} \frac{f_i(tH_iu)}{t} = C_i(u), \tag{13}$$

then the following statements are true:

(i) $C_i(\alpha u) = |\alpha| \, C_i(u)$ and $C_i(u_1 + u_2) \leq C_i(u_1) + C_i(u_2)$.

(ii) Define the function $h_i(u, v, t) : \mathbb{R}^n \times \mathbb{R}^{m_i} \times \mathbb{R} \mapsto \mathbb{R}$,

$$h_i(u, v, t) \triangleq \frac{1}{t} \left[f_i(v + tH_i u) - f_i(v) \right]. \tag{14}$$

Then the following pointwise limit holds:

$$\lim_{t \to \infty} h_i(u, v, t) = C_i(u). \tag{15}$$

Moreover, the convergence is uniform on any compact set of (u, v).

(iii) For any v and u, we have that

$$f_i(v + H_i u) - f_i(v) \leq C_i(u). \tag{16}$$

Remark 3: Intuitively speaking, one can interpret f_i as a potential field and the derivative of f_i as the force generated by sensor i (if it is differentiable). By (16), we know that the force from the potential field f_i along the u direction cannot exceed $C_i(u)$ (or $C_i(u)/\|u\|$ to normalize). On the other hand, Equation (15) implies that this bound is achievable.

We now give the sufficient condition for the robustness of the estimator.

Theorem 1 (Sufficient condition): If the following conditions hold:

1) $C_i(u)$ is well defined for all $u \in \mathbb{R}^n$ and all $i \in \mathcal{S}$;
2) the following inequality holds for all non-zero u:

$$\sum_{i \in \mathcal{I}} C_i(u) < \sum_{i \in \mathcal{I}^c} C_i(u), \ \forall \mathcal{I} \in \mathbb{C}, \tag{17}$$

then the estimator g is robust.

Our goal is to prove that there exists a $\beta(z)$, such that for any $t \geq \beta(z)$, $\|u\| = 1$, $a \in \mathcal{A}$, the following inequality holds:

$$\sum_{i \in \mathcal{S}} f_i(y_i - H_i \times tu) < \sum_{i \in \mathcal{S}} f_i(y_i - H_i \times (t+1)u). \tag{18}$$

As a result, any point $\|\hat{x}\| \geq \beta(z) + 1$ cannot be the solution of the optimization problem since there exists a better point $(\|\hat{x}\| - 1)\hat{x}/\|\hat{x}\|$. Therefore, we must have $\|g(y)\| \leq \beta(z) + 1$ and hence the estimator is robust.

Suppose the set of malicious sensors is $\mathcal{I}$, to prove (18), we will first look at benign sensors. Due to the uniform convergence of $h_i(u, v, t)$ to $C_i(u)$ on $\mathcal{U} \times \{-z_i\}$ shown in Lemma 2, given any $\delta > 0$ we can always find a finite constant N_i depending on δ and z_i such that for all $t \geq N_i(\delta, z_i)$, the following inequality holds:

$$h_i(-z_i, u, t) = \frac{1}{t}\left[f_i(tH_i u - z_i) - f_i(-z_i)\right] \geq C_i(u) - \delta, \tag{19}$$

for any $\|u\| = 1$. By (12), we can derive that

$$f_i((t+1)H_i u - z_i) - f_i(tH_i u - z_i) \geq C_i(u) - \delta. \tag{20}$$

We define $\beta(z) \triangleq \max_{1\leq i\leq m} N_i(\delta, z_i)$ and fix δ to be

$$\delta = \frac{1}{m}\min_{\|u\|=1}\min_{\mathcal{I}\in\mathbb{C}}\left(\sum_{i\in\mathcal{I}^c} C_i(u) - \sum_{i\in\mathcal{I}} C_i(u)\right). \tag{21}$$

Hence, for $i = 1, \ldots, m$, if $t > \beta_\delta(z)$ we have

$$\begin{aligned} &f_i((t+1)H_i u - z_i) - f_i(tH_i u - z_i) \\ &\qquad\qquad\qquad\qquad \geq C_i(u) - \delta,\ \forall\|u\| = 1. \end{aligned} \tag{22}$$

Since for good sensors, $z_i = y_i$, we know that

$$\begin{aligned} &\sum_{i\in\mathcal{I}^c}\left[f_i((t+1)H_i u - z_i) - f_i(tH_i u - z_i)\right] \\ &\qquad \geq \sum_{i\in\mathcal{I}^c} C_i(u) - (m-p)\delta,\ \forall\|u\| = 1. \end{aligned} \tag{23}$$

We now consider malicious sensors. By Lemma 2 (iii), we know that for $i \in \mathcal{I}$, and any u

$$\sum_{i\in\mathcal{I}} f_i(y_i - tH_i u) - \sum_{i\in\mathcal{I}} f_i(y_i - (t+1)H_i u) \leq \sum_{i\in\mathcal{I}} C_i(-u). \tag{24}$$

Hence from (21), (24) and (23), we know that

$$\begin{aligned} &\sum_{i\in\mathcal{S}} f_i(y_i - (t+1)H_i u) - \sum_{i\in\mathcal{S}} f_i(y_i - tH_i u) \\ &\qquad \geq \sum_{i\in\mathcal{I}^c} C_i(u) - \sum_{i\in\mathcal{I}} C_i(u) - (m-p)\delta > 0, \end{aligned}$$

which proves (18).

Remark 4: Assuming that y_i is a scalar and $w = 0$, Fawzi et al. [16] prove that the state can be exactly recovered under the integrity attack if and only if for all $u \neq 0$, there are at least $2p+1$ non-zero $H_i u$. Notice that if for some $u \neq 0$, there are less than $2p+1$ non-zero $H_i u$, then we can choose $\mathcal{I}$ to contain the largest p $H_i u$ and thus violate (17). As a result, our sufficient condition is stronger than the ones proposed in [16]. The main reason is that we seek to use convex optimization to solve the state estimation problem, while in [16], a combinatorial optimization problem is needed to recover the state.

We next give necessary conditions for the robustness of the estimator.

Theorem 2 (Necessary Condition I): If $C_i(u)$ is well defined for all $u \in \mathbb{R}^n$ and all $i \in \mathcal{S}$ but there exist some $\|u_0\| = 1,\ \mathcal{I}_0 \in \mathbb{C}$ such that

$$\sum_{i \in \mathcal{I}_0} C_i(u_0) > \sum_{i \in \mathcal{I}_0^c} C_i(u_0), \tag{25}$$

then the estimator is not robust to the attack.

The robustness of the estimator is equivalent to that the optimal estimate $\hat{x}$ satisfies $\|\hat{x}\| \leq \mu(z)$ for all $a \in \mathcal{A}$, where μ is a real-valued function. To this end, we will prove that for any $r > 0$, there exists a y such that all $\hat{x}$ that satisfies $\|\hat{x}\| \leq r$ cannot be the optimal solution of (5).

We will first look at the compromised sensors. For every $\delta > 0$ we can always find a finite constant $N_i(\delta)$ such that for any $\hat{x} \in \{\hat{x} : \|\hat{x}\| \leq r\}$ and for all $t > N_i$, the following inequality holds:

$$\begin{aligned} & f_i(tH_iu_0 - H_i\hat{x}) - f_i(tH_iu_0 - H_i(\hat{x}+u_0)) \\ & f_i((t+1)H_iu_0 - H_i(\hat{x}+u_0)) - f_i(tH_iu_0 - H_i(\hat{x}+u_0)) \\ \geq & h_i(u_0, -H_i(\hat{x}+u_0), t) \geq C_i(u_0) - \delta,\ \forall i \in \mathcal{I}_0. \end{aligned} \tag{26}$$

The first inequality is derived from (12). The second inequality is due to the uniform convergence of $h_i(u, v, t)$ to $C_i(u)$ on $\{u_0\} \times \{v : v = -H_ix + u_0,\ \|x\| \leq r\}$.

Let us choose

$$\delta = \frac{1}{m}\left(\sum_{i \in \mathcal{I}_0} C_i(u_0) - \sum_{i \in \mathcal{I}_0^c} C_i(u_0)\right),$$

and $t = \max_{i \in \mathcal{I}_0} N_i(\delta)$ and $y_i = tH_iu_0$ for all $i \in \mathcal{I}_0$, then we know for any $\|\hat{x}\| \leq r$,

$$\sum_{i \in \mathcal{I}_0} \left[f_i(y_i - H_i\hat{x}) - f_i(y_i - H_i(\hat{x}+u_0))\right]$$

$$\geq \sum_{i \in \mathcal{I}_0} C_i(u_0) - p\delta.$$

Now let us look at the benign sensors. By Lemma 2 (iii) we have

$$f_i(z_i - H_i(\hat{x} + u_0)) - f_i(z_i - H_i\hat{x}) \\ \leq C_i(u_0),\ \forall i \in \mathcal{I}\backslash\mathcal{I}_{m_0}. \quad (27)$$

From (26) and (27),

$$\sum_{i\in\mathcal{S}} f_i(y_i - H_i(\hat{x} + u_0)) - \sum_{i\in\mathcal{S}} f_i(y_i - H_i\hat{x}) \\ \leq \sum_{i\in\mathcal{I}_0^c} C_i(u_0) - \sum_{i\in\mathcal{I}_0} C_i(u_0) + p\delta < 0.$$

Thus for such a y satisfying

$$y_i = \begin{cases} z_i, & \text{if } i \in \mathcal{I}_0^c \\ tH_iu_0, & \text{if } i \in \mathcal{I}_0, \end{cases}$$

$\hat{x} + u_0$ is a better estimate than all $\hat{x}$ satisfying $\|\hat{x}\| \leq r$. Since r is an arbitrary positive real number, we can conclude that the estimator is not robust.

Theorem 3 (Necessary Condition II): If there exists $u_0 \in \mathbb{R}^n$ and $j \in \mathcal{I}$ such that

$$\lim_{t\to\infty} \frac{f_i(tH_iu_0)}{t} \to +\infty, \quad (28)$$

then the estimator is not robust to the attack.

Before proving Theorem 3, we need the following lemma whose proof is reported in appendix.

Lemma 3: If the condition (28) holds, for any $M > 0$ and for all v in a compact set $\mathcal{V} \subset \mathbb{R}^{m_i}$, there exists N (depending on M and the set $\mathcal{V}$) such that the following inequality holds:

$$h_j(u_0, v, t) > M, \forall v \in \mathcal{V} \quad (29)$$

Now we are ready to prove the theorem.

Similar to Theorem 2, we will prove that for any $r > 0$, there exists a y such that all $\hat{x}$ that satisfies $\|\hat{x}\| \leq r$ cannot be the optimal solution of (5).

We first look at any sensor i, where $i \neq j$. Since a continuous function achieves its supremum on a compact set, we know that the following supremum is well defined (not infinite)

$$\sup_{\|\hat{x}\|\leq r} [f(z_i - H_i(\hat{x} + u_0)) - f(z_i - H_i\hat{x})] = M_i,$$

which implies that for all $\|\hat{x}\| \leq r$, we can find $M > 0$, such that

$$\sum_{i \neq j} f(z_i - H_i(\hat{x} + u_0)) - \sum_{i \neq j} f(z_i - H_i\hat{x}) \leq M. \tag{30}$$

Now let us consider sensor j. Due to Lemma 3, we can find a t, such that for all $\|\hat{x}\| \leq r$, the following inequality holds:

$$h_j(u_0, -H_j(\hat{x} + u_0), t) > M.$$

Using Lemma 1, we have

$$\begin{aligned}
&f((t+1)H_j u_0 - H_j(\hat{x} + u_0)) - f(tH_j u_0 - H_j(\hat{x} + u_0)) \\
&= f(tH_j u_0 - H_j\hat{x}) - f(tH_j u_0 - H_j(\hat{x} + u_0)) \\
&\geq h_j(u_0, -H_j(\hat{x} + u_0), t) > M.
\end{aligned} \tag{31}$$

Now consider the following y

$$y_i = \begin{cases} z_i, & \text{if } i \neq j \\ tH_j u_0, & \text{if } i = j, \end{cases}$$

Combining (30) and (31), we know that for all $\|\hat{x}\| \leq r$, the following inequality holds

$$\sum_{i \in \mathcal{S}} f(y_i - H_i(\hat{x} + u_0)) - \sum_{i \in \mathcal{S}} f(y_i - H_i\hat{x}) < M - M = 0,$$

which implies that the optimal solution of (5) cannot be inside the ball $\{\hat{x} : \|\hat{x}\| \leq r\}$. Now since $r > 0$ is arbitrary, we know the estimator is not robust. Before continuing on, we would like to provide some remarks on the main result. First, it is worth noticing that the existence of a well defined limit of $f_i(tH_i u)/t$ is crucial for the robustness of g as Theorem 3 suggested. For example, the least square estimator cannot be robust since f_i is in quadratic form. Using the potential field and force analogies in Remark 3, one can interpret the results presented in this section as: the estimator g is robust if the force generated by any sensor is bounded and if the combined force of any collection of p sensors is no greater than the combined force of the remaining $m - p$ sensors.

Secondly, one can see that the conditions proved in Theorem 1, 2 and 3 are very tight, with only a trivial gap where the LHS of (25) equals the RHS.

Finally, we want to point out that the condition (17) is non-trivial to check since it requires us to verify against all possible u. In the next subsection, we consider a special case where each y_i

is a scalar and provide a more conservative but verifiable sufficient condition for the robustness of the estimator.

IV. Scalar Measurement Case: More Analysis

In this section, we specialize our results to the scalar measurement case, *i.e.*, $m_i = 1,\ \forall i \in \mathcal{S}$. Throughout this section, we assume that the following limit is well-defined:

$$\alpha_i \triangleq \lim_{t \to \infty} f_i(t)/t. \tag{32}$$

It is not difficult to prove that $C_i(u) = |\alpha_i H_i u|$. With slight abuse of notation, define $C_i \triangleq \alpha_i H_i$, then $C_i(u) = |C_i u|$. For any index set $\mathcal{I} = \{i_1, \dots, i_l\} \subset \mathcal{S}$, define

$$C_{\mathcal{I}} \triangleq \begin{bmatrix} C_{i_1} \\ \vdots \\ C_{i_l} \end{bmatrix}. \tag{33}$$

From Theorem 1 and Theorem 2, we have the following sufficient and necessary conditions for robustness of g.

Proposition 1:

(a) If for all possible index set $\mathcal{I}$ and all non-zero $u \in \mathbb{R}^n$ the following inequality holds:

$$\|C_{\mathcal{I}} u\|_1 = \sum_{i \in \mathcal{I}} |C_i u| < \sum_{i \in \mathcal{I}^c} |C_i u| = \|C_{\mathcal{I}^c} u\|_1, \tag{34}$$

then the estimator g is robust.

(b) If there exists an index set $\mathcal{I}$ and a $u \in \mathbb{R}^n$ such that the following inequality holds:

$$\|C_{\mathcal{I}} u\|_1 > \|C_{\mathcal{I}^c} u\|_1, \tag{35}$$

then the estimator g is not robust.

The main difficulty here is to validate (34) for all non-zero u. In the next theorem, we can find a more conservative but more practically useful sufficient condition for the robustness, by eliminating u from (34).

Theorem 4: If for any index set $\mathcal{I} \subset \mathcal{S}$ with cardinality p, the optimal value of the following optimization problem is strictly less than 1:

$$\begin{aligned} &\underset{K \in \mathbb{R}^{n \times (m-p)}}{\text{minimize}} && \|C_{\mathcal{I}} K\|_1 \\ &\text{subject to} && K C_{\mathcal{I}^c} = \mathrm{I}_n, \end{aligned} \tag{36}$$

then the estimator g is robust.

Let $K \in \mathbb{R}^{n\times(m-p)}$ such that $KC_{\mathcal{I}^c} = \mathrm{I}_n$. Denote $\xi = C_{\mathcal{I}^c}u$. We have $C_{\mathcal{I}}u = C_{\mathcal{I}}K\xi$. Therefore, if for all $\xi \neq 0$, $\|C_{\mathcal{I}}K\xi\|_1 < \|\xi\|_1$, *i.e.*, $\|C_{\mathcal{I}}K\|_1 < 1$, then

$$\|C_{\mathcal{I}}u\|_1 < \|C_{\mathcal{I}^c}u\|_1.$$

By enumerating all possible $\mathcal{I}$ we can conclude the proof.

Notice that (36) is not necessary. Since ξ is in the column space of $C_{\mathcal{I}^c}$, ξ may not be able to take all possible value in $\mathbb{R}^{m-p}$.

Similarly, we can find a more practically useful version for the necessary condition implied by Theorem 2. By enumerating all $(C_{\mathcal{I}}, C_{\mathcal{I}^c})$ and utilizing the following result, we can identify whether g is robust for a given H or not.

Theorem 5: If there exists an index set $\mathcal{I}$ such that the following inequality holds:

$$\|C_{\mathcal{I}}C_{\mathcal{I}^c}^{+}\|_1 > (\sqrt{m-p}+1)/2, \tag{37}$$

where $C_{\mathcal{I}^c}^{+}$ is the Moore-Penrose pseudo inverse of $C_{\mathcal{I}^c}$, then the estimator g is not robust.

The following lemma, whose proof is given in the appendix, is needed for the proof of Theorem 5:

Lemma 4: Let $\xi \in \mathbb{R}^m$ such that $\xi = \xi_{\|} + \xi_{\perp}$, where $\xi_{\|}$ and $\xi_{\perp}$ are perpendicular to each other. Then the following inequality holds:

$$\|\xi_{\|}\|_1 \leq \frac{\sqrt{m}+1}{2}\|\xi\|_1. \tag{38}$$

Moreover, the above inequality is achievable when

$$\xi = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \; \xi_{\|} = \frac{1}{2}\begin{bmatrix} 1+m^{-1/2} \\ m^{-1/2} \\ \vdots \\ m^{-1/2} \end{bmatrix}, \; \xi_{\perp} = \frac{1}{2}\begin{bmatrix} 1-m^{-1/2} \\ -m^{-1/2} \\ \vdots \\ -m^{-1/2} \end{bmatrix}.$$

We are now ready to prove Theorem 5: To prove g is not robust, from Proposition 1 we only need to show there exists a u such that $\|C_{\mathcal{I}}u\|_1 > \|C_{\mathcal{I}^c}u\|_1$ if (37) holds. Since $\left\|C_{\mathcal{I}}C_{\mathcal{I}^c}^{+}\right\|_1 > (\sqrt{m-p}+1)/2$, we can find $\xi \in \mathbb{R}^{m-p}$, such that

$$\left\|C_{\mathcal{I}}C_{\mathcal{I}^c}^{+}\xi\right\|_1 > \frac{\sqrt{m-p}+1}{2}\left\|\xi\right\|_1.$$

Now we can decompose $\xi = \xi_{\|} + \xi_{\perp}$, where $\xi_{\|}$ belongs to the column space of $C_{\mathcal{I}^c}$ and $\xi_{\perp}$ is perpendicular to the column space of $C_{\mathcal{I}^c}$. By the property of Moore-Penrose inverse, $C_{\mathcal{I}^c}^{+}\xi_{\perp} = 0$. Therefore,

$$\left\|C_{\mathcal{I}} C_{\mathcal{I}^c}^{+} \xi\right\|_1 = \left\|C_{\mathcal{I}} C_{\mathcal{I}^c}^{+} \xi_{\|}\right\|_1 .$$

On the other hand, since $\xi \in \mathbb{R}^{m-p}$, by Lemma 4, we have

$$\frac{\sqrt{m-p}+1}{2} \left\|\xi\right\|_1 \geq \left\|\xi_{\|}\right\|_1 ,$$

which implies that

$$\left\|C_{\mathcal{I}} C_{\mathcal{I}^c}^{+} \xi_{\|}\right\|_1 > \left\|\xi_{\|}\right\|_1 .$$

Since $\xi_{\|}$ belongs to the column space of $C_{\mathcal{I}^c}$, there exists a u, such that $C_{\mathcal{I}^c}u = \xi_{\|}$. Therefore, we can find a u, such that

$$\left\|C_{\mathcal{I}} u\right\|_1 > \left\|C_{\mathcal{I}^c} u\right\|_1 ,$$

which completes the proof.

V. Concluding Remarks

We have studied the robust estimation problem where p out of m sensors are under attack. The malicious measurements can be arbitrarily manipulated and thus a robust estimator which can give a reliable estimate is needed. Our interest is not to study any concrete estimator in presence of attacks. Instead, we have considered a general class of estimators which integrate a large number of important estimators as special cases and given sufficient and necessary conditions for the robustness of the estimator. Moreover, we have presented more analytical results in the scalar measurement case to render the sufficient and necessary conditions more ready to use. Future works include the robustness analysis for the dynamical state estimation problem.

VI. Appendix

Proof of Lemma 2:

(i) If $\alpha = 0$, then clearly $C_i(0) = 0$. On the other hand, if $\alpha \neq 0$, from the definition in (13), we have

$$\begin{aligned} C_i(\alpha u) &= \lim_{t\to\infty} \frac{1}{t} f_i(|\alpha| t H_i u) \\ &= |\alpha| \lim_{t\to\infty} \frac{1}{|\alpha|\, t} f_i(|\alpha| t H_i u) = |\alpha|\, C_i(u). \end{aligned}$$

Due to the scaling property of $C_i(u)$ and the convexity of f_i, we have

$$C_i(u_1+u_2) = 2C_i\left(\frac{u_1+u_2}{2}\right) \leq C_i(u_1)+C_i(u_2).$$

Therefore, we know that C_i is actually a semi-norm on $\mathbb{R}^n$

(ii) Based on the convexity of f_i, we obtain

$$2f_i(\frac{tH_iu}{2}) \leq f_i(v+tH_iu)+f_i(-v), \tag{39}$$

$$f_i(tH_iu) \geq 2f_i(\frac{2v+tH_iu}{2}) - f(2v). \tag{40}$$

Dividing both sides of (39) and (40) by t and taking limit over t, we have

$$C_i(u) \leq \liminf_{t\to\infty} \frac{1}{t} f_i(v+tH_iu) + \lim_{t\to\infty}\frac{1}{t}f_i(-v), \tag{41}$$

$$C_i(u) \geq \limsup_{t\to\infty} \frac{2}{t} f_i(v+\frac{t}{2}H_iu) - \lim_{t\to\infty}\frac{1}{t}f_i(2v). \tag{42}$$

Since $\lim_{t\to\infty} f_i(-v)/t = \lim_{t\to\infty} f_i(2v)/t = 0$, from (42) and (41) we have the following pointwise limit

$$\lim_{t\to\infty} h_i(u,v,t) = C_i(u).$$

Notice that for a fixed (u,v), by Lemma 1, $h(u,v,t)$ is monotonically non-decreasing with respect to t. Furthermore, $C_i(u)$ is continuous since it is a semi-norm. Therefore, by Dini's theorem [31], $h(u,v,t)$ converges uniformly to $C_i(u)$ on a compact set of (u,v).

(iii) By Lemma 1, we have

$$f_i(v+H_iu) - f_i(v) = f_i(H_iu)$$
$$\leq \lim_{t\to} \frac{f_i(tH_iu)}{t} = C_i(u).$$

Proof of Lemma 3:

From (28) and (39), it is easy to see that $h_j(u_0,v,t)$ diverges to infinity for all v, *i.e.*,

$$h_j(u_0,v,t) \to +\infty. \tag{43}$$

Next we will show this divergence is also uniform. Denote $\mathcal{W}_t \triangleq \{v : h_j(u_0,v,t) > M\}$. Since $h_j(u_0,v,t)$ is continuous (f_i is continuous due to convexity), each $\mathcal{W}_t$ is open. By Lemma 1,

$h_j(u_0, v, t)$ is monotonically non-decreasing in t. Therefore, $\mathcal{W}_t \subseteq \mathcal{W}_{t'}$ if $t \leq t'$. Since for each v there exists t such that $h_j(u_0, v, t) > M$ from (43),

$$\bigcup_{t\geq 0} \mathcal{W}_t = \mathbb{R}^{m_i}.$$

Therefore, the collection $\{\mathcal{W}_t\}$ is an open cover for the compact subset $\mathcal{V}$. Thus, we can find a finite cover $\mathcal{W}_{t_1}, \ldots, \mathcal{W}_{t_l}$ that covers $\mathcal{V}$, i.e.,

$$\mathcal{V} \subseteq \mathcal{W}_{t_1} \cup \mathcal{W}_{t_2} \cup \cdots \cup \mathcal{W}_{t_l}. \tag{44}$$

Now we can define $N = \max(t_1, \ldots, t_l)$. Since $\mathcal{W}_t$ is non-decreasing with respect to t, the RHS of (44) is $\mathcal{W}_N$. For any $t \geq N$, we have

$$\mathcal{V} \subseteq \mathcal{W}_N \subseteq \mathcal{W}_t,$$

which combined with the definition of $\mathcal{W}_t$ finishes the proof.

Proof of Lemma 4:

Geometrically, $\xi_{||}$ can be written as $\xi_{||} = \xi/2 + r$, where $r \in \{r : \|r\|_2 = \|\xi\|_2/2\}$. As a result, we have

$$\begin{aligned}\|\xi_{||}\|_1 \leq \frac{1}{2}\|\xi\|_1 + \|r\|_1 \leq \frac{1}{2}\|\xi\|_1 + \sqrt{m}\|r\|_2 \\ = \frac{1}{2}\|\xi\|_1 + \frac{\sqrt{m}}{2}\|\xi\|_2 \leq \frac{1}{2}\|\xi\|_1 + \frac{\sqrt{m}}{2}\|\xi\|_1,\end{aligned}$$

The first inequality is due to the triangle inequality of any norm. The second and third inequalities are due to the fact that for an m dimensional vector ξ,

$$\|\xi\|_2 \leq \|\xi\|_1 \leq \sqrt{m}\|\xi\|_2.$$

Without loss of generality, let us assume that $\|\xi\|_1 = 1$. The achievability of (38) is easy to verify.

REFERENCES

[1] S. Massoud Amin and B. Wollenberg, "Toward a smart grid: power delivery for the 21st century," *IEEE Power and Energy Mag.*, vol. 3, no. 5, pp. 34–41, Sep. 2005.

[2] S. A. Boyer, "SCADA: supervisory control and data acquisition," *Instrument Engineers' Handbook, Volume Three: Process Software and Digital Networks*, p. 357, 2002.

[3] T. M. Chen, "Stuxnet, the real start of cyber warfare? [editor's note]," *IEEE Network*, vol. 24, no. 6, pp. 2–3, 2010.

[4] D. P. Fidler, "Was stuxnet an act of war? decoding a cyberattack," *IEEE Security & Privacy*, vol. 9, no. 4, pp. 56–59, 2011.

[5] A. A. Cárdenas, S. Amin, and S. Sastry, "Research challenges for the security of control systems," in *Proc. Conf. Hot Topics in Security*. Berkeley, CA, USA: USENIX Association, 2008, pp. 1–6.

[6] H. P. Lopuhaa and P. J. Rousseeuw, "Breakdown points of affine equivariant estimators of multivariate location and covariance matrices," *The Annals of Statistics*, vol. 19, no. 1, pp. 229–248, 1991.

[7] E. Handschin, F. Schweppe, J. Kohlas, and A. Fiechter, "Bad data analysis for power system state estimation," *IEEE Trans. Power Apparatus and Systems*, vol. 94, no. 2, pp. 329–337, Mar. 1975.

[8] L. Mili, T. Van Cutsem, and M. Ribbens-Pavella, "Bad data identification methods in power system state estimation - a comparative study," *IEEE Power Engineering Review*, vol. PER-5, no. 11, pp. 27–28, Nov. 1985.

[9] H. Sandberg, A. Teixeira, and K. H. Johansson, "On security indices for state estimators in power networks," in *First Workshop on Secure Control Systems*, 2010.

[10] L. Xie, Y. Mo, and B. Sinopoli, "Integrity data attacks in power market operations," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 659–666, 2011.

[11] Y. Liu, M. Reiter, and P. Ning, "False data injection attacks against state estimation in electric power grids," in *Proc. ACM Conf. Computer and Commun. Security*, 2009.

[12] J. Kim, L. Tong, and R. J. Thomas, "Data framing attack on state estimation," *IEEE J. Selected Areas in Commun.*, vol. 32, no. 7, pp. 1460–1470, Jul. 2014.

[13] F. Pasqualetti, A. Bicchi, and F. Bullo, "Consensus computation in unreliable networks: a system theoretic approach," *IEEE Transactions on Automatic Control*, vol. 57, no. 1, pp. 90–104, Jan 2010.

[14] F. Pasqualetti, F. Dorfler, and F. Bullo, "Cyber-physical attacks in power networks: models, fundamental limitations and monitor design," in *Proc. IEEE Conf. Decision and Control and European Control Conf.*, 2011, pp. 2195–2201.

[15] S. Sundaram, M. Pajic, C. Hadjicostis, R. Mangharam, and G. J. Pappas, "The wireless control network: monitoring for malicious behavior," in *Proc. IEEE Conf. Decision and Contro*, 2010.

[16] H. Fawzi, P. Tabuada, and S. Diggavi, "Security for control systems under sensor and actuator attacks," in *Proc. IEEE Conf. Decision and Control*, 2012, pp. 3412–3417.

[17] M. S. Chong, M. Wakaiki, and J. P. Hespanha, "Observability of linear systems under adversarial attacks," in *Proc. IEEE Amer. Control conf.*, 2015.

[18] M. Pajic, J. Weimer, N. Bezzo, P. Tabuada, O. Sokolsky, I. Lee, and G. J. Pappas, "Robustness of attack-resilient state estimators," in *Proc. ACM/IEEE Int. Conf. Cyber-Physical Systems*, Apr. 2014, pp. 163–174.

[19] Y. Mo, E. Garone, A. Casavola, and B. Sinopoli, "False data injection attacks against state estimation in wireless sensor networks," in *Proc. IEEE Conf. Decision and Control*, 2010, pp. 5967–5972.

[20] Y. Mo and B. Sinopoli, "False data injection attacks in cyber physical systems," in *First Workshop on Secure Control Systems*, 2010.

[21] S. Kassam, H. V. Poor *et al.*, "Robust techniques for signal processing: A survey," *Proc. IEEE*, vol. 73, no. 3, pp. 433–481, 1985.

[22] R. A. Maronna, D. R. Martin, and V. J. Yohai, *Robust Statistics: Theory and Methods*. NJ: Wiley, 2006.

[23] P. J. Huber and E. M. Ronchetti, *Robust Statistics*. NJ:Wiely, 2009.

[24] F. R. Hampel, “A general qualitative definition of robustness,” *The Annals of Mathematical Statistics*, vol. 42, no. 6, pp. 1887–1896, 1971.

[25] D. L. Donoho and P. J. Huber, “The notion of breakdown point,” *A Festschrift for Erich L. Lehmann*, pp. 157–184, 1983.

[26] F. R. Hampel, “The influence curve and its role in robust estimation,” *J. the American Statistical Association*, vol. 69, no. 346, pp. 383–393, 1974.

[27] V. J. Yohai and R. H. Zamar, “High breakdown-point estimates of regression by means of the minimization of an efficient scale,” *J. the American Statistical Association*, 2012.

[28] P. Rousseeuw and C. Croux, “Explicit scale estimators with high breakdown point,” 1992.

[29] O. Hössjer, “Rank-based estimates in the linear model with high breakdown point,” *J. the American Statistical Association*, vol. 89, no. 425, pp. 149–158, 1994.

[30] P. J. Rousseeuw, “Least median of squares regression,” *J. the American Statistical Association*, 2012.

[31] W. Rudin, *Principles of mathematical analysis*. McGraw-Hill New York, 1964, vol. 3.