Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnline CONFERENCE PAPER ARCHIVE

IFAC PapersOnLine 51-23 (2018) 289–294

Secure Static State Estimation: A Large Deviation Approach

Xiaoqiang Ren * **Yilin Mo** ** **Karl H. Johansson** *

* *ACCESS Linnaeus Center, School of Electrical Engineering, KTH Royal Institute of Technology, 114 28 Stockholm, Sweden, (e-mails: xiaren, kallej@kth.se)*

** *School of Electrical and Electronics Engineering, Nanyang Technological University, Singapore, (e-mail: ylmo@ntu.edu.sg)*

Abstract: This paper studies static state estimation based on measurements from a set of sensors, a subset of which can be compromised by an attacker. The measurements from a compromised sensor can be manipulated arbitrarily by the adversary. A new notion is adopted to indicate the performance of an estimator, that is, the asymptotic exponential rate, with which the worst-case probability of estimate lying outside certain ball centered at the true underlying state goes to zero. An optimal estimator, which computes Chebyshev centers and only utilizes the information contained in the averaged measurements, is proposed. Numerical examples are given to elaborate the results.

Keywords: Secure, State estimation, Byzantine sensors, Large deviation

1. INTRODUCTION

Background: Static state estimation has a wide range of applications in power system Schweppe and Handschin (1974). In a power system, numerous sensors, which usually possess limited capacity, are spatially deployed and connected via ubiquitous wireless/wired communication networks. This makes it nearly impossible to guarantee the security of every single sensor or communication channel. Therefore, security problems in a power system, or a networked system in general, has attracted much attention recently Mo et al. (2012); Teixeira et al. (2015).

Our Work, its Contributions and Related Literature: Robust estimation has been studied over decades to deal with the uncertainties of input data Hampel (1974); Kassam and Poor (1985); Huber (2011). The robustness is usually measured by influence functions or breakdown point Huber (2011), and several celebrated estimators have been developed, such as M-estimator, L-estimator, and R-estimator. The limitation of robustness theory is the assumption that the bad data are independent Huber (2011). However, in our work, the fact that compromised sensors may cooperate and the estimation is done sequentially makes the "bad" data correlated.

Recently, dynamic state estimation with possibly Byzantine sensors has attracted much attention. Most of approaches in existing literature can be classified into two categories: stacked measurements Fawzi et al. (2014); Pajic et al. (2017); Mishra et al. (2017) and Kalman filter decomposition Mo and Garone (2016); Liu et al. (2017). Fawzi et al. (2014) used the stacked measurements from time k to $k+T-1$ to estimate the state at time k and provided l_0 and l_1-based state estimation procedures. Since deterministic systems are concerned, the l_0-based procedure can exactly recover the state. Pajic et al. (2017) extended deterministic systems in Fawzi et al. (2014) to ones with bounded measurement noises and obtained upper bounds of estimation error of both l_0 and l_1-based estimators. Mishra et al. (2017) further studied stochastic systems with unbounded noises and proposed a notion of ϵ-effective attack. The state estimation there is in essence an attack detection problem, i.e., a Chi-squared test is applied to the residues and the standard Kalman filter output based on the measurements from the largest set of sensors that are deemed ϵ-effective attack-free is used as the state estimate. Notice that to detect the ϵ-effective attack-free sensors correctly with high probability, the window size T must be large enough. Mo and Garone (2016); Liu et al. (2017) used local estimators at each sensor and proposed a LASSO based fusion scheme. However, their approach imposes some strong constraints on the system dynamics. Furthermore, the estimate error of the proposed algorithm when there are indeed attacks is not specifically characterized.

In this paper, we deal with scenarios where noises are not necessarily bounded and give a different characterization of the estimator performance, i.e., the decaying rate of the worst-case probability that the estimation error is larger than some value δ rather than the worst-case error in Pajic et al. (2017); Mo and Garone (2016); Liu et al. (2017) and estimation error covariance in Mishra et al. (2017). This is partially motivated by the following three observations. Firstly, if unbounded noises are involved, the worst-case estimation error might result in too conservative system designs. Notice also that even for the bounded noises cases studied in Pajic et al. (2017), the upper bound of the worst-case estimation error thereof increases with

⋆ The work by X. Ren and K. H. Johansson was supported by the Knut and Alice Wallenberg Foundation, the Swedish Foundation for Strategic Research, and the Swedish Research Council.

10.1016/j.ifacol.2018.12.050

respect to the window size T, which counters intuition since more information should lead to better estimation accuracy. Secondly, to mitigate the bad effects caused by possible Byzantine sensors, one has to accumulate much enough information, that is, the time window size T in the aforementioned stacked measurements approach should be large enough. In this case, the decaying rate of a probability is able to characterize its value well enough. Lastly, the system operator may pre-define the error threshold δ according to the performance specification, which leads to a more flexible system design.

In the subsequent sections, we focus on the problem of secure state estimation with Byzantine sensors. A fusion center aims to estimate a vector state $x \in \mathbb{R}^n$ from measurements collected by m sensors, among which n sensors might be compromised. Without imposing any restrictions on the attacker's capabilities, we assume that the compromised sensors can send arbitrary messages. This work, which deals with static state estimation though, provides insights into the dynamic state estimation we are planning to investigate in the future. We note that secure static state estimation with Byzantine sensors was studied in Han et al. (2015) as well, which, however, focused on resilience analysis about a generic convex optimization based estimator and used the worst-case estimate error as the performance metric. What is more, Han et al. (2015) studied the one-shot scenario, while in this work the observations are taken sequentially, the possible temporal correlations make the analysis more challenging.

The main contributions of this work are summarized as follows.

(1) We propose a new metric to characterize the performance of an estimator when an adversary may be present. This metric is reasonable and a flexible system design can be achieved based on the *a priori* performance requirement.
(2) An optimal estimator is proposed, which computes Chebyshev centers based on the averaged measurements (Theorem 4).

Notations: $\mathbb{R}$ ($\mathbb{R}_+$) is the set of (nonnegative) real numbers. $\mathbb{N}$ ($\mathbb{N}_+$) is the set of nonnegative (positive) integers. For a vector $x \in \mathbb{R}^n$, define $\|x\|_0$ as the "zero norm", i.e., the number of nonzero elements of the vector x. For a vector $x \in \mathbb{R}^n$, the support of x, denoted by $\mathrm{supp}(x)$, is the set of indices of nonzero elements:

$$\mathrm{supp}(x) \triangleq \{i \in \{1, 2, \ldots, n\} : x_i \neq 0\}.$$

Define $\mathbf{1}$ as the column vector, of which each element is one and the size should be clear from the context. For a matrix $\boldsymbol{M} \in \mathbb{R}^{m\times n}$, unless stated otherwise, $\boldsymbol{M}_i$ represents the i-th row of $\boldsymbol{M}$, and $\boldsymbol{M}_{\mathcal{I}}$ the matrix obtained from $\boldsymbol{M}$ after removing all of the rows except those indexed by $\mathcal{I}$, i,e, $\boldsymbol{M}_i$'s with $i \in \mathcal{I}$.

2. PROBLEM FORMULATION

2.1 System Model

Consider the problem of estimating the state $x \in \mathbb{R}^n$ using m sensors' measurements. The measurement equation for sensor i is

$$z_i(k) = H_i x + w_i(k),$$

where $z_i(k) \in \mathbb{R}$ is the real-valued "true" measurement collected by the sensor i at time $k \in \mathbb{N}_+$, $H_i \in \mathbb{R}^{1\times n}$ is the output matrix associated with sensor i, $w_i(k) \in \mathbb{R}$ is the observation noise. It is assumed that $w_i(k)$ is Gaussian distributed with zero mean and covariance $\mathbb{E}[(w_i(k))^2] = W_i$ for any i, k. Furthermore, $w_i(k)$ are independent across the sensors and over time, i.e., $\mathbb{E}[w_{i_1}(k_1)w_{i_2}(k_2)] = 0$ if $i_1 \neq i_2$ or $k_1 \neq k_2$.

In the presence of attacks, the measurement equation may be

$$y_i(k) = z_i + a_i(k),$$

where $y_i(k) \in \mathbb{R}$ is the manipulated measurement that arrives at the fusion center and $a_i(k) \in \mathbb{R}$ is the bias injected by the attacker.

Let $\mathcal{M} \triangleq \{1, \ldots, m\}$ be the index set of all the sensors. We assume the attacks are q-sparse:

Assumption 1. (q-sparse attack). There exists an index set $\mathcal{C} \subset \mathcal{M}$ such that

(1) for any sensor $i \in \mathcal{M} \setminus \mathcal{C}$, $a_i(k) = 0$ for any time k.
(2) $|\mathcal{C}| = q$.

The sparse attack model, which is conventional in the related literature Fawzi et al. (2014); Pajic et al. (2017); Mishra et al. (2017); Mo and Garone (2016); Liu et al. (2017); Han et al. (2015); Ren et al. (2018), says that the set of compromised sensors is somewhat "constant" over time. This is in essence the only restriction we impose on the attacker's capability, since the bias $a_i(k)$ of a compromised sensor may take any value.

To introduce measurement knowledge of the attacker, we need the following definitions. Define the measurement from all sensors at time k to be a column vector:

$$\boldsymbol{y}(k) \triangleq [y_1(k),\ y_2(k),\ \cdots\ y_m(k)]^\top \in \mathbb{R}^m. \tag{1}$$

We further define $\boldsymbol{Y}(k)$ as a matrix of all measurements from time 1 to time k:

$$\boldsymbol{Y}(k) \triangleq [\boldsymbol{y}(1),\ \boldsymbol{y}(2),\ \cdots\ \boldsymbol{y}(k)] \in \mathbb{R}^{m\times k}. \tag{2}$$

The quantities $\boldsymbol{a}(k), \boldsymbol{A}(k)$ are defined in the same manner.

Assumption 2. (Attacker's knowledge). The adversary is assumed to know the system parameters (i.e., each H_i and W_i) and true state, has the causal knowledge of observations from the compromised sensors, and owns unlimited memory, i.e., $\boldsymbol{Y}(k)_{\mathcal{C}}$ and $\boldsymbol{A}(k)$ are known to the attacker at time k.

The above assumption is prevailing in the related literature as well, see Mishra et al. (2017); Mo and Garone (2016); Liu et al. (2017); Han et al. (2015); Ren et al. (2018). By the knowledge about sensors, an attacker can develop the parameters H_i and W_i. To obtain the true state, the attacker may deploy its own sensor network. Though it might be difficult to obtain the accurate parameters and true state for an attacker in practice, this assumption is de facto when dealing with potential worst-case attacks. We should note that this assumption is in accordance with the Kerckhoffs's principle Shannon (1949), namely the security of a system should not rely on its obscurity.

With Assumptions 1 and 2, the measurements from compromised sensors might take any value and might be correlated across sensors and over time.

We assume that the system knows the number q, but does not know the exact set of compromised sensors $\mathcal{C}$. The quantity q might be determined by the *a priori* knowledge about the quality of each sensor. Alternatively, the quantity q may be viewed as a design parameter, which indicates the resilience level that the system is willing to introduce.

Let $\boldsymbol{H} = [H_1^\top, H_2^\top, \dots, H_m^\top]^\top$ be the measurement matrix. We assume that the matrix $\boldsymbol{H}$ is $2q$-observable:

Assumption 3. The measurement matrix $\boldsymbol{H}$ is $2q$-observable, i.e., for every set $\mathcal{I} \subset \mathcal{M}$ with $|\mathcal{I}| = m - 2q$, the matrix $\boldsymbol{H}_\mathcal{I}$ is observable.

It has been shown in Fawzi et al. (2014) that the $2q$-observability of measurement matrix is the necessary and sufficient condition to recover the exact state under q-sparse attacks when there are no observation noises.

2.2 Performance Metric

At time k, given measurements from all the sensors $\boldsymbol{Y}(k)$, the fusion center generates a state estimate $\hat{x}_k$. Notice that the estimator, denoted by f_k, might be random, i.e., given $\boldsymbol{Y}(k)$, $\hat{x}_k$ is random variable governed by certain probability measure on $\mathbb{R}^n$ that is determined by f_k. Let the system strategy $f \triangleq (f_1, f_2, \dots)$ be a sequence of estimators from time 1 to ∞.

Similarly, at time k, given the measurements $\boldsymbol{Y}(k)$, the bias $\boldsymbol{A}(k-1)$, the set of compromised sensors $\mathcal{C}$ and true state x, the bias $\boldsymbol{a}(k)$ is generated according to some probability measure on $\mathbb{R}^m$. This bias injection mechanism at time k is denoted by g_k and the attack strategy by $g \triangleq (g_1, g_2, \dots)$. Let $\mathcal{G}$ be the set of all attack strategies such that the generated bias $\boldsymbol{a}(k)$ satisfies the q-sparse attack model in Assumption 1.

In this paper, we are concerned with the worst-case scenario. Given a system strategy f, we define

$$e(f,k,\delta) \triangleq \sup_{\mathcal{C}\subset\mathcal{M}, g\in\mathcal{G}, x\in\mathbb{R}^n} \mathbb{P}_{f,g,x,\mathcal{C}}\left(\|\hat{x}_k - x\|_2 > \delta\right) \quad (3)$$

as the worst-case probability that the distance between the estimate at time k and true state is larger than certain value $\delta \in \mathbb{R}_+$ considering all possible attack strategy, the set of compromised sensors and the true state. We use $\mathbb{P}_{f,g,x,\mathcal{C}}$ to emphasize that the probability depends on the estimator f, attack strategy g, the true state x, and the set of compromised sensors $\mathcal{C}$.

Ideally, one wants to design an estimator f such that $e(f,k,\delta)$ is minimized at any time k for any δ. However, it is quite difficult to analyze $e(f,k,\delta)$ when k takes finite values since computing the probability of error usually involves numerical integration. Therefore, we consider a asymptotic estimation performance, i.e., the exponential rate that the worst-case probability goes to zero:

$$r(f,\delta) \triangleq \liminf_{k\to\infty} -\frac{\log e(f,k,\delta)}{k}. \quad (4)$$

Obviously, for any δ, the system would like to maximize $r(f,\delta)$ by designing the estimator f.

2.3 Problems of Interest

What is the optimal estimator f in the sense that the rate $r(f,\delta)$ is maximized for a given δ?

3. OPTIMAL ESTIMATOR IN THE PRESENCE OF ATTACKS

In this section, we provide an estimator as Chebyshev centers and sketch out the proof of its optimality due to the page limitation.

To proceed, we first present in Theorem 1 a preliminary results on a random variable, which is then utilized to develop the "nice" structure of an optimal estimator shown in Corollary 1. The latter lays the foundations of proving the optimality of our proposed estimator.

3.1 Preliminaries

We first need the following lemma:

Lemma 1. Let $\mathcal{A}_1, \dots, \mathcal{A}_n$ be a finite collection of events with the same underlying sample space, i.e., $\mathbb{P}(\cup_{j=1}^n \mathcal{A}_j) \le 1$. Then it holds that

$$\mathbb{P}(\cap_{j=1}^n \mathcal{A}_j) \ge \sum_{j=1}^n \mathbb{P}(\mathcal{A}_j) - n + 1. \quad (5)$$

Proof. If $n = 1$, it trivially holds.

When $n \ge 2$, we prove this by induction. Notice that

$$\begin{aligned}\mathbb{P}(\cap_{j=1}^2 \mathcal{A}_j) &= \sum_{j=1}^2 \mathbb{P}(\mathcal{A}_j) - \mathbb{P}(\cup_{j=1}^2 \mathcal{A}_j)\\ &\ge \sum_{j=1}^2 \mathbb{P}(\mathcal{A}_j) - 1,\end{aligned}$$

which concludes the case when $n = 2$. Suppose (5) holds for some $n-1$ with $n \ge 3$, then

$$\begin{aligned}\mathbb{P}(\cap_{j=1}^n \mathcal{A}_j) =& \mathbb{P}(\cap_{j=1}^{n-1} \mathcal{A}_j) + \mathbb{P}(\mathcal{A}_n) - \mathbb{P}((\cap_{j=1}^{n-1} \mathcal{A}_j) \cup \mathcal{A}_j)\\ &\ge \sum_{j=1}^{n-1} \mathbb{P}(\mathcal{A}_j) - n + 2 + \mathbb{P}(\mathcal{A}_n) - 1\\ &\ge \sum_{j=1}^{n} \mathbb{P}(\mathcal{A}_j) - n + 1.\end{aligned}$$

The proof is thus complete. □

Let $\mathcal{B}_\delta(x)$ denote the closed ball center at $x \in \mathbb{R}^n$ with radius $\delta > 0$:

$$\mathcal{B}_\delta(x) \triangleq \{y \in \mathbb{R}^n : \|y - x\|_2 \le \delta\}.$$

Theorem 1. Given any random variable $y \in \mathbb{R}^n$, there always exist $x^* \in \mathbb{R}^n$ such that $\mathbb{P}(y \notin \mathcal{B}_\delta(x)) \ge 1/(n+1)$ holds for every $x \notin \mathcal{B}_\delta(x^*)$.

Proof. Let $\mathcal{A}$ denote the set of x such that random variable y has a high probability lying in its neighborhood:

$$\mathcal{A} \triangleq \{x : \mathbb{P}(y \in \mathcal{B}_\delta(x)) > n/(n+1))\}. \quad (6)$$

Then it suffices to show that there exists $x^* \in \mathbb{R}^n$ such that

$$\mathcal{A} \subseteq \mathcal{B}_\delta(x^*). \quad (7)$$

In the following argument, the dimension n is fixed.

If $\mathcal{A}$ only contains $j \leq (n+1)$ elements, say, $x_1, \ldots, x_j$. Then Lemma 1 together with (6) yields that

$$\mathbb{P}(y \in \cap_{i=1}^{j} \mathcal{B}_\delta(x_i)) > 0,$$

which means that the set $\cap_{i=1}^{j} \mathcal{B}_\delta(x_i)$ is not empty. Then $\mathcal{A} \subseteq \mathcal{B}_\delta(x^*)$ for any $x^* \in \cap_{i=1}^{j} \mathcal{B}_\delta(x_i)$.

If $\mathcal{A}$ contains $j > (n+1)$ elements (j might be infinite). Then again by Lemma 1, one obtains that for *any* $n+1$ elements, say, $x_1, \ldots, x_{n+1}$, there holds

$$\mathbb{P}(y \in \cap_{i=1}^{n+1} \mathcal{B}_\delta(x_i)) > 0,$$

that is, $\cap_{i=1}^{n+1} \mathcal{B}_\delta(x_i) \neq \emptyset$. Since $\mathcal{B}_\delta(x)$ is compact and convex for any x, then Helly's theorem Danzer and Klee (1963) means that

$$\cap_{x \in \mathcal{A}} \mathcal{B}_\delta(x) \neq \emptyset.$$

Then $\mathcal{A} \subseteq \mathcal{B}_\delta(x^*)$ for any $x^* \in \cap_{x \in \mathcal{A}} \mathcal{B}_\delta(x)$. The proof is thus complete. □

3.2 Compressed and Deterministic Estimator

Notice that a generic estimator f_k might *randomly* generate the estimate $\hat{x}_k$ based on *all* the information contained in $\boldsymbol{Y}(k)$. In other words, given a different $\boldsymbol{Y}(k)$, the probability measure that governs $\hat{x}_k$ might be different. In this subsection, however, we shall show that without loss of optimality, one may only consider estimators with certain "nice" structure.

To proceed, we define an operator $\operatorname{avg}(\cdot)$ that averages each row of the inputed real-valued matrix, i.e., for any matrix $\boldsymbol{M} \in \mathbb{R}^{n_1 \times n_2}$,

$$\operatorname{avg}(\boldsymbol{M}) \triangleq \boldsymbol{M}\mathbf{1}/n_2,$$

where $\mathbf{1}$ is the column vector with each element being 1. Then one can see that $\operatorname{avg}(\boldsymbol{Y}(k))$ is a vector in $\mathbb{R}^m$ and the i-th element is the empirical mean of the observation from time 1 to k available for sensor i.

We use $\mathbb{P}_f(\hat{x}_k | \boldsymbol{Y}(k))$ to denote the conditional probability measure of estimate $\hat{x}_k$ given any system strategy $f \in \mathcal{F}$ and the information $\boldsymbol{Y}(k)$ [1].

Definition 1. A system strategy f is said to be compressed if it only utilizes the averaged information $\operatorname{avg}(\mathbf{Y}(k))$ to generate estimate $\hat{x}_k$ at each time k, that is, the conditional probability measures satisfy

$$\mathbb{P}_f(\hat{x}_k \in \mathcal{A} | \boldsymbol{Y}(k)) = \mathbb{P}_f(\hat{x}_k \in \mathcal{A} | \boldsymbol{Y}'(k)) \tag{8}$$

for any Borel set $\mathcal{A} \subset \mathbb{R}^n$ whenever $\operatorname{avg}(\boldsymbol{Y}(k)) = \operatorname{avg}(\boldsymbol{Y}'(k))$.

Let $\mathcal{F}$ ($\mathcal{F}_{\rm c}$, resp.) be the set of all possible (compressed, resp.) system strategies. In the following theorem, we show that $\mathcal{F}$ and $\mathcal{F}_{\rm c}$ are equivalent in our setting.

Theorem 2. For any $f \in \mathcal{F}$, there exists another estimator $f' \in \mathcal{F}_{\rm c}$ such that

$$e(f', k, \delta) \leq e(f, k, \delta), \quad \forall \delta > 0, k.$$

Proof. Due to the page limitation, we only present the critical part. The details can be found in the journal version. The proof is constructive: for any $f \in \mathcal{F}$, we let f' satisfy (9) and (10). For any $y \in \mathbb{R}^m$, Borel set $\mathcal{A} \subset \mathbb{R}^n$, and time k,

$$\begin{aligned} &\mathbb{P}_{f'}(\hat{x}_k \in \mathcal{A} | \operatorname{avg}(\boldsymbol{Y}(k)) = y) \\ &= \int_{\mathbb{R}^{m \times k}} \mathbb{P}_f(\hat{x}_k \in \mathcal{A} | \boldsymbol{Y}(k) = Y) \\ &\qquad \mathrm{d}\mathbb{P}(\boldsymbol{Z}(k) = Y | \operatorname{avg}(\boldsymbol{Z}(k)) = y), \end{aligned} \tag{9}$$

where $\mathrm{d}\mathbb{P}(\boldsymbol{Z}(k) = Y | \operatorname{avg}(\boldsymbol{Z}(k)) = y)$ is the derivative of conditional probability measure $\mathbb{P}(\boldsymbol{Z}(k) | \operatorname{avg}(\boldsymbol{Z}(k)) = y)$ at the point $\boldsymbol{Z}(k) = Y$. Notice that the conditional probability measure $\mathbb{P}(\boldsymbol{Z}(k) | \operatorname{avg}(\boldsymbol{Z}(k)) = y)$ is well-defined since $\operatorname{avg}(\boldsymbol{Z}(k))$ is a sufficient statistic of the "true" measurements $\boldsymbol{Z}(k)$ for the underlying state x, i.e., for any state x,

$$\mathbb{P}_x(\boldsymbol{Z}(k) | \operatorname{avg}(\boldsymbol{Z}(k)) = y) = \mathbb{P}(\boldsymbol{Z}(k) | \operatorname{avg}(\boldsymbol{Z}(k)) = y).$$

Furthermore, $f' \in \mathcal{F}_{\rm c}$ means that

$$\mathbb{P}_{f'}(\hat{x}_k \in \mathcal{A} | \boldsymbol{Y}(k)) = \mathbb{P}_{f'}(\hat{x}_k \in \mathcal{A} | \boldsymbol{Y}'(k)) \tag{10}$$

for any Borel set $\mathcal{A} \subset \mathbb{R}^n$ whenever $\operatorname{avg}(\boldsymbol{Y}(k)) = \operatorname{avg}(\boldsymbol{Y}'(k))$. □

Remark 1. Intuitively, only measurements from benign sensors provide "useful information" needed to estimate the underlying state, while the most harmful compromised sensors will merely generate disturbing noises. In our case, the averaged information $\operatorname{avg}(\mathbf{Y}(k))$ can *fully* summarize the information contained in measurements from benign sensors due to the fact that $\operatorname{avg}(\mathbf{Y}(k))$ is a sufficient statistic when there is no attacker. Therefore, one may only consider a compressed estimator that only utilizes the averaged information each time.

Definition 2. A system strategy f is said to be compressed and deterministic if there exists one possibly time-dependent function $\tilde{f}_k : \mathbb{R}^m \to \mathbb{R}^n$ such that the estimate at each time k

$$\hat{x}_k = \tilde{f}_k(\operatorname{avg}(\boldsymbol{Y}(k))).$$

By the above definition, under a compressed and deterministic system strategy: 1) only the "averaged information" contained in $\boldsymbol{Y}(k)$ (i.e., $\operatorname{avg}(\boldsymbol{Y}(k))$) is utilized. 2) given $\operatorname{avg}(\boldsymbol{Y}(k))$, the estimate is deterministic (rather than random) since $\tilde{f}$ is a function.

Let $\mathcal{F}_{\rm cd}$ be the set of all compressed and deterministic system strategies. Then one sees that $\mathcal{F}_{\rm cd} \subset \mathcal{F}_{\rm c} \subset \mathcal{F}$. In the following theorem, we show that $\mathcal{F}_{\rm c}$ and $\mathcal{F}_{\rm cd}$ are equivalent in our setting.

Theorem 3. For any $f \in \mathcal{F}_{\rm c}$, there exists another estimator $f' \in \mathcal{F}_{\rm cd}$ such that

$$r(f', \delta) \geq r(f, \delta), \quad \forall \delta > 0.$$

Proof. Notice that a compressed strategy $f \in \mathcal{F}_{\rm c}$ can be completely characterized by the sequence of conditional probability measures from time 1 to ∞: $(\mathbb{P}_f(\hat{x}_1 | \operatorname{avg}(\boldsymbol{Y}(1))), \mathbb{P}_f(\hat{x}_2 | \operatorname{avg}(\boldsymbol{Y}(2))), \ldots)$. The estimator f' is given as follows: for any $y \in \mathbb{R}^m$ and time k,

$$\mathbb{P}_{f'}(\mathcal{A} | \operatorname{avg}(\boldsymbol{Y}(k)) = y) = \begin{cases} 1 & \text{when } x^*(k) \in \mathcal{A}, \\ 0 & \text{otherwise,} \end{cases}$$

where $x^*(k)$ is such that $\mathbb{P}_f(x \notin \mathcal{B}_\delta(\hat{x}_k)) \geq 1/(n+1)$ holds for every $x \notin \mathcal{B}_\delta(x^*(k))$, the existence of which is guaranteed by Theorem 1. Then one obtains that for every $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$:

[1] Notice that if the estimator is deterministic at time k, then the conditional probability measure degenerates and only takes value on one point in $\mathbb{R}^n$.

$$\mathbb{P}_f\left(\hat{x}_k \notin \mathcal{B}_\delta(x)\middle|\, \mathrm{avg}(\boldsymbol{Y}(k)) = y\right)$$
$$\geq \mathbb{P}_{f'}\left(\hat{x}_k \notin \mathcal{B}_\delta(x)\middle|\, \mathrm{avg}(\boldsymbol{Y}(k)) = y\right)/(n+1). \quad (11)$$

Then one obtains that for every time k and $\delta > 0$:

$$e(f,k,\delta) \geq e(f',k,\delta)/(n+1). \quad (12)$$

Recall that $e(f,k,\delta)$ is the worst-case probability defined in (3). Then it follows that for any $\delta > 0$:

$$\begin{aligned} r(f,\delta) &= \liminf_{k\to\infty} -\frac{\log e(f,k,\delta)}{k} \\ &\leq \liminf_{k\to\infty} -\frac{\log e(f',k,\delta)/(n+1)}{k} \\ &= \liminf_{k\to\infty} -\frac{\log e(f',k,\delta)}{k} \\ &= r(f',\delta). \end{aligned} \quad (13)$$

The proof is thus complete. □

With the above two theorems, the following can be readily obtained:

Corollary 1. Without loss of optimality, one can only consider a compressed and deterministic system strategy in $\mathcal{F}_{\mathrm{cd}}$, i.e., for any $f \in \mathcal{F}$, there exists another estimator $f' \in \mathcal{F}_{\mathrm{sd}}$ such that

$$r(f',\delta) \geq r(f,\delta), \quad \forall \delta > 0.$$

3.3 Optimal Estimator as Chebyshev Centers

To proceed, we need the following definitions. The distance of a point $x_0 \in \mathbb{R}^n$ to a bounded and non-empty set $\mathcal{A} \subset \mathbb{R}^n$ is defined as

$$\mathrm{dist}(x_0,\mathcal{A}) = \sup\{\|x - x_0\|_2 : x \in \mathcal{A}\}.$$

For a bounded and non-empty set $\mathcal{A} \subseteq \mathbb{R}^n$, its radius $\mathrm{rad}(\mathcal{A}) \in \mathbb{R}_+$ and Chebyshev center $c(\mathcal{A}) \in \mathbb{R}^n$ are computed by

$$\mathrm{rad}(\mathcal{A}) = \min_{x_0\in\mathbb{R}^n} \mathrm{dist}(x_0,\mathcal{A}),$$
$$c(\mathcal{A}) = \arg\min_{x_0\in\mathbb{R}^n} \mathrm{dist}(x_0,\mathcal{A}).$$

Given $y \in \mathbb{R}^m, x \in \mathbb{R}^n$, define its "distance" $d(y,x)$ as the optimal value of the following optimization problem:

$$\begin{aligned} \underset{a\in\mathbb{R}^m}{\text{minimize}} \quad & \frac{1}{2}\sum_{i=1}^{m}(y_i - H_i x + a_i)^2/W_i \\ \text{subject to} \quad & \|a\|_0 = q. \end{aligned} \quad (14)$$

Further define the set $\mathcal{X}(y,\phi), \phi \geq 0$ as the set of x such that the distance to y is upper bounded by ϕ, i.e.,

$$\mathcal{X}(y,\phi) \triangleq \{x \in \mathbb{R}^n : d(y,x) \leq \phi\}.$$

Given $\delta \geq 0$, define $\mathbb{X}(y,\delta)$ as the set of x such that the radius of $\mathcal{X}(y,\phi)$ is upper bounded by δ:

$$\mathbb{X}(y,\delta) \triangleq \bigcup_{\phi\in\Phi(y,\delta)} \mathcal{X}(y,\phi), \quad (15)$$

where $\Phi(y,\delta)$ is given by

$$\Phi(y,\delta) \triangleq \{\phi \geq 0 : \mathrm{rad}(\mathcal{X}(y,\phi)) \leq \delta\}.$$

Let f^* be the estimator such that the estimate at time k is the Chebyshev center of $\mathbb{X}(\mathrm{avg}(\boldsymbol{Y}(k)),\delta)$, i.e.,

$$f_k^*(\boldsymbol{Y}(k)) = c\left(\mathbb{X}(\mathrm{avg}(\boldsymbol{Y}(k)),\delta)\right). \quad (16)$$

The optimality of f^* is shown in the following theorem.

Theorem 4. Given any $\delta > 0$, the estimator f^* is optimal in the sense that it maximizes the rate in (4), i.e., for any admissible estimator $f \in \mathcal{F}$, there holds

$$r(f^*,\delta) \geq r(f,\delta).$$

Proof. We first provide the upper bound of $r(f,\delta)$ for any $f \in \mathcal{F}$. This leverages Corollary 1 and show that any $f \in \mathcal{F}_{\mathrm{sd}}$ possessing performance $r(f,\delta)$ larger than the upper bound would require certain set of $\mathrm{avg}(\boldsymbol{Y}(k))$ to be in two different non-intersecting balls, which is impossible. We then show that our proposed estimator f^* can achieve this upper bound. The performance is analyzed mainly using the large deviation theory.

4. NUMERICAL SIMULATIONS

Theorem 1 plays a critical role in the derivation of the main results in this paper, in particular, Theorem 3. It is, however, somewhat abstract. Therefore, in this section we illustrate Theorem 1 by several examples when the dimension $n = 2$.

In all the examples, δ is fixed to be 1. We let y has different distributions and plot the region $\{x : \mathbb{P}(y \notin \mathcal{B}_\delta(x)) < 1/3\}$. Both Figs. 1 and 2 verify Theorem 1 since in each case, the region $\{x : \mathbb{P}(y \notin \mathcal{B}_\delta(x)) < 1/3\}$ can be covered by a disk with radius not larger than 1.

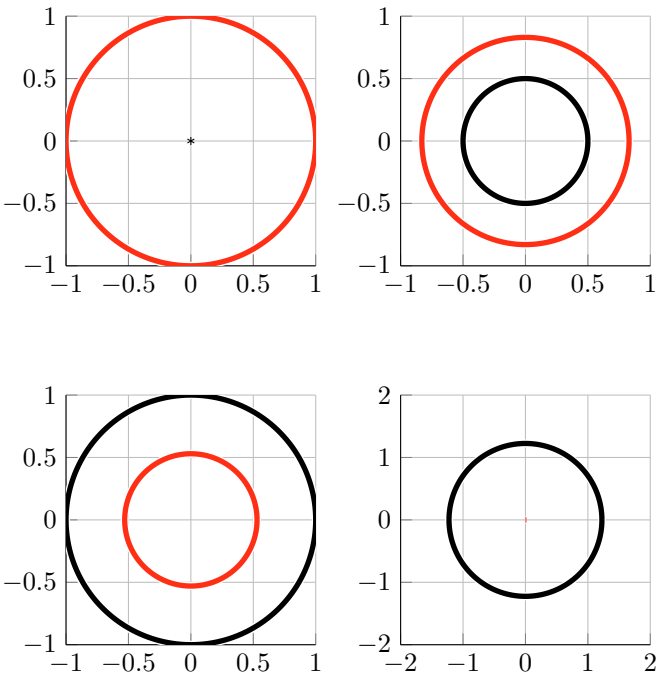

Fig. 1. The random variable y is at point $[0,0]$ with probability 1 in the top-left sub-figure, and is uniformly distributed in the disk with boundary being the black circle in the other three ones. Every black circle is centered at the origin and is with radius being $0.5, 1, \sqrt{3/2}$, respectively. The region $\{x : \mathbb{P}(y \notin \mathcal{B}_\delta(x)) < 1/3\}$ is empty in the down-right sub-figure, and is in the disk with circumference being the red circle in the other three ones.

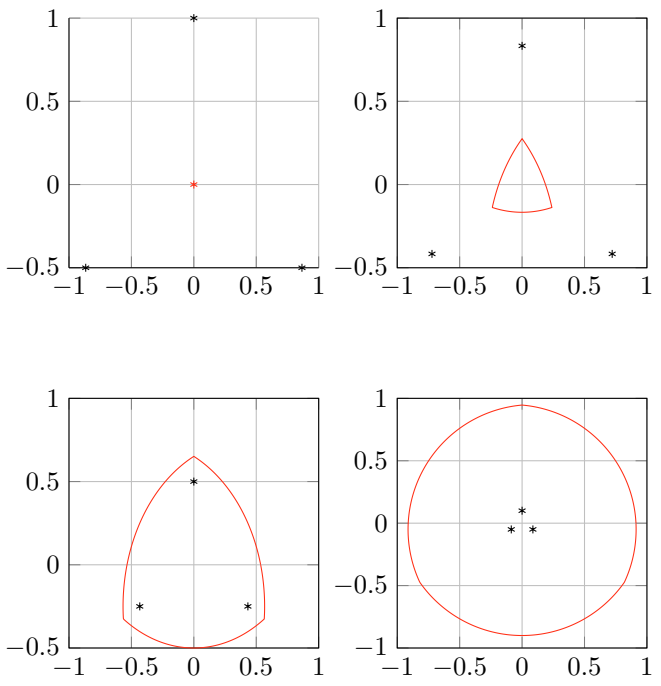

Fig. 2. In all the four cases, the random variable y may be one of the three points (indicated by black marks) uniformly with probability 1/3. The three points are vertices of an equilateral triangle, of which the circumcenter is the origin and circumradius is $1, 5/6, 0.5, 0.1$, respectively. The region $\{x : \mathbb{P}(y \notin \mathcal{B}_\delta(x)) < 1/3\}$ is a singleton (i.e., the origin) in the top-left sub-figure, and is the area inside the red circle in the other three ones.

5. CONCLUSION

In this paper, we studied secure static state estimation problem with Byzantine sensors. A new metric was proposed to characterize the performance of an estimator, i.e., the decaying rate of the worst-case probability that the estimation error is larger than certain value. An optimal estimator, which only utilizes the averaged measurements and computes Chebyshev centers, was given.

REFERENCES

Danzer, L. and Klee, V. (1963). *Helly's theorem and its relatives*. American Mathematical Society Providence, RI.

Fawzi, H., Tabuada, P., and Diggavi, S. (2014). Secure estimation and control for cyber-physical systems under adversarial attacks. *IEEE Transactions on Automatic Control*, 59(6), 1454–1467.

Hampel, F.R. (1974). The influence curve and its role in robust estimation. *Journal of the American Statistical Association*, 69(346), 383–393.

Han, D., Mo, Y., and Xie, L. (2015). Convex optimization based state estimation against sparse integrity attacks. *CoRR*, abs/1511.07218. URL `http://arxiv.org/abs/1511.07218`.

Huber, P.J. (2011). *Robust statistics*. Springer.

Kassam, S.A. and Poor, H.V. (1985). Robust techniques for signal processing: A survey. *Proceedings of the IEEE*, 73(3), 433–481.

Liu, X., Mo, Y., and Garone, E. (2017). Secure dynamic state estimation by decomposing kalman filter. *IFAC-PapersOnLine*, 50(1), 7351–7356.

Mishra, S., Shoukry, Y., Karamchandani, N., Diggavi, S.N., and Tabuada, P. (2017). Secure state estimation against sensor attacks in the presence of noise. *IEEE Transactions on Control of Network Systems*, 4(1), 49–59.

Mo, Y. and Garone, E. (2016). Secure dynamic state estimation via local estimators. In *IEEE 55th Conference on Decision and Control (CDC)*, 5073–5078. IEEE.

Mo, Y., Kim, T.H.J., Brancik, K., Dickinson, D., Lee, H., Perrig, A., and Sinopoli, B. (2012). Cyber–physical security of a smart grid infrastructure. *Proceedings of the IEEE*, 100(1), 195–209.

Pajic, M., Lee, I., and Pappas, G.J. (2017). Attack-resilient state estimation for noisy dynamical systems. *IEEE Transactions on Control of Network Systems*, 4(1), 82–92.

Ren, X., Yan, J., and Mo, Y. (2018). Binary hypothesis testing with Byzantine sensors: Fundamental trade-off between security and efficiency. *IEEE Transactions on Signal Processing*, 66(6), 1454–1468.

Schweppe, F.C. and Handschin, E.J. (1974). Static state estimation in electric power systems. *Proceedings of the IEEE*, 62(7), 972–982.

Shannon, C.E. (1949). Communication theory of secrecy systems. *Bell Labs Technical Journal*, 28(4), 656–715.

Teixeira, A., Sou, K.C., Sandberg, H., and Johansson, K.H. (2015). Secure control systems: A quantitative risk management approach. *IEEE Control Systems*, 35(1), 24–45.