European Journal of Control 74 (2023) 100837

Contents lists available at ScienceDirect

European Journal of Control

journal homepage: www.elsevier.com/locate/ejcon

Resilient set-based state estimation for linear time-invariant systems using zonotopes

Muhammad Umar B. Niazi [a,b,*], Amr Alanwar [c], Michelle S. Chong [d], Karl H. Johansson [b]

[a] Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[b] Division of Decision and Control Systems, Digital Futures, School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Stockholm, Sweden
[c] School of Computer Science and Engineering, Constructor University, Germany
[d] Control Systems Technology Section, Department of Mechanical Engineering, Eindhoven University of Technology, the Netherlands

ARTICLE INFO

Article history:
Received 4 May 2023
Accepted 8 June 2023
Available online 15 June 2023

Recommended by Prof. T Parisini

Keywords:
Resilient estimation
Set-based methods
Zonotopic filtering

ABSTRACT

This paper considers the problem of set-based state estimation for linear time-invariant (LTI) systems under time-varying sensor attacks. Provided that the LTI system is stable and observable via every single sensor and that at least one sensor is uncompromised, we guarantee that the true state is always contained in the estimated set. We use zonotopes to represent these sets for computational efficiency. However, we show that intelligently designed stealthy attacks may cause exponential growth in the algorithm's worst-case complexity. We present several strategies to handle this complexity issue and illustrate our resilient zonotope-based state estimation algorithm on a rotating target system.

1. Introduction

Cyber physical systems are quite vulnerable to attackers who can maliciously manipulate their sensor measurements and cause undesirable disruptions in the system. To ensure that the corrupted sensor data does not degrade estimation accuracy, and thereby guaranteeing reasonable control performance, several *resilient* or *secure* state estimation techniques have been proposed (c.f. Alanwar et al. [1], Chong et al. [11], He et al. [13], Kim et al. [15], Pajic et al. [16], Shoukry et al. [21]). In the presence of additive sensor attacks, these techniques ensure estimation accuracy up to a certain neighborhood of the true state, modulo noise and disturbances, where the estimation error bound remains uninfluenced by the attacker. However, the error bounds of these estimators are quite conservative, and it is difficult to obtain precise robust guarantees.

For the state estimation of dynamical systems, stochastic filtering approaches, such as Kalman filters, assume that the statistics of the underlying random process generating the process and measurement noise is known. In applications where noise statistics cannot be known, such filtering approaches perform rather poorly [18]. For non-stochastic uncertainties, $\mathcal{H}_\infty$ filters and observers provide a robust solution to the state estimation problem; however, they turn out to be overly conservative [22]. To evade these limitations and at the same time obtain precise robust guarantees, the set-based zonotopic filtering paradigm has proven to be very promising [6,7,12], with many real-world applications including fault diagnosis in industrial systems [9], underwater robotics [14], vehicle localization [10], and leakage detection in water distribution networks [17].

In safety-critical applications, guaranteed state inclusion in a bounded set is crucial to provably avoid unsafe regions in the state space. This motivates the need for set-based state estimation to obtain a set of all possible states under unknown disturbances and measurement errors belonging to known bounded sets. In this regard, set-based estimation has a long history and was first studied by [8] in 1971. Recent works on the topic are [6,7] and the references therein.

To the best of our knowledge, the literature on resilient set-based state estimation when some of the sensors are vulnerable to adversarial attacks is scarce. Only [19] and its journal version [20] present a set-based resilient state estimation technique that relies on reachability. However, for obtaining an accurate set-based estimate, they require that the full state vector can be measured by any subset of sensors with cardinality equal to the number of safe sensors, which is quite a restrictive assumption. Without this, their guarantees on the estimation accuracy become very conservative.

* Corresponding author at: Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.
E-mail addresses: niazi@mit.edu (M.U.B. Niazi), aalanwar@constructor.university (A. Alanwar), m.s.t.chong@tue.nl (M.S. Chong), kallej@kth.se (K.H. Johansson).

https://doi.org/10.1016/j.ejcon.2023.100837

In this paper, we do not require the full state vector to be measured by any subset of sensors, but the LTI system is observable via every sensor. This assumption is required since we do not limit the number of attacked sensors to be less than half the number of sensors, and we also allow the attacker to change the set of compromised sensors at any time. Subject to these assumptions, we present a zonotope-based state estimation algorithm for LTI systems under sensor attacks. We guarantee that the true state is always included in the estimated set. A strength of our proposed scheme is that we can handle attacks compromising different sensors over time as long as at least one sensor remains untouched. Further, if the attacker compromises the same set of attacked sensors over time, we provide a detection scheme to identify the set of attacked sensors. One major drawback of the zonotope-based algorithm is its complexity, which can increase exponentially in the worst-case if stealthy attacks are employed. We discuss and demonstrate complexity reduction schemes to help with the implementation of our zonotope-based resilient state estimation algorithm. All used data and code to recreate our findings are publicly available.[1]

The rest of the paper is organized as follows. We define the notations used and background needed in Section 2. Next, the problem and assumptions are stated in Section 3. We then present the resilient zonotopic state estimation algorithm and guarantee that the true state is always within the estimated set in Section 4. In Section 5, we present cases where our results can be sharpened and discuss the complexity of the proposed scheme. The efficacy of the algorithm is illustrated by an example in Section 6. We conclude the paper with Section 7.

2. Notations and preliminaries

2.1. Notations

The set of real numbers and integers are denoted by $\mathbb{R}$ and $\mathbb{Z}$, respectively, and $\mathbb{Z}_{\geq i} \doteq \{i, i+1, i+2, \ldots\}$. A finite set of integers $\{i, i+1, i+2, \ldots, i+k\}$ is denoted as $\mathbb{Z}_{[i,i+k]}$. The Euclidean norm of a vector $x \in \mathbb{R}^n$ is denoted as $\|x\| \doteq \sqrt{x^T x}$ and the maximum norm as $\|x\|_\infty \doteq \max_{i\in\{1,\ldots,n\}} |x_i|$. Given a signal $v : \mathbb{Z}_{\geq 0} \to \mathbb{R}^n$, we denote its restriction to the domain $[0, k]$ by $v_{[0,k]}$, for some $k \in \mathbb{Z}_{\geq 0}$. For a set $\mathcal{S}$, $|\mathcal{S}|$ denotes its cardinality. Given sets $\mathcal{S}_1, \ldots, \mathcal{S}_n$, we denote their collection as $\mathcal{S} = \{\mathcal{S}_i\}_{i\in\mathbb{Z}_{[1,n]}}$.

2.2. Set representations

Given a center $c_z \in \mathbb{R}^n$ and generator matrix $G_z \in \mathbb{R}^{n\times\xi_z}$, a *zonotope* $\mathcal{Z} \subset \mathbb{R}^n$ is the set

$$\mathcal{Z} \doteq \{c_z + G_z\beta_z : \beta_z \in [-1, 1]^{\xi_z}\}$$

where ξ_z is the number of generators of $\mathcal{Z}$. Since a zonotope can be completely characterized by its center and generator matrix, the notation $\mathcal{Z} = \langle c_z, G_z\rangle$ is used throughout the paper for brevity.

A matrix $L \in \mathbb{R}^{n'\times n}$ multiplied with a zonotope $\mathcal{Z}$ yields $L\mathcal{Z} = \langle Lc_z, LG_z\rangle$. Given two zonotopes $\mathcal{Z}_1 = \langle c_{z_1}, G_{z_1}\rangle$ and $\mathcal{Z}_2 = \langle c_{z_2}, G_{z_2}\rangle$, each being a subset of $\mathbb{R}^n$, their Minkowski sum is given by

$$\mathcal{Z}_1 \oplus \mathcal{Z}_2 = \langle c_{z_1} + c_{z_2}, [G_{z_1} \quad G_{z_2}]\rangle.$$

A *constrained zonotope* is given by

$$\mathcal{Z} \doteq \{c_z + G_z\beta_z : \beta_z \in [-1, 1]^{\xi_z}, A\beta_z = b\}$$

where $A \in \mathbb{R}^{n\times\xi_z}$ and $b \in \mathbb{R}^n$ with $n \in \mathbb{Z}_{>0}$. The radius of a zonotope, or a constrained zonotope, is given by

$$\mathrm{rad}(\mathcal{Z}) = \min \Delta \text{ subject to } \mathcal{Z} \subset \Delta\mathbb{B}^n(c_z)$$

i.e., the radius Δ of a minimal n-dimensional Euclidean ball $\mathbb{B}^n(c_z)$ centered at c_z and inscribing $\mathcal{Z}$.

To summarize, a zonotope is an affine transformation of a hypercube and a constrained zonotope is an affine transformation of linearly constrained hypercube.

3. Problem definition

Consider an LTI system in discrete-time

$$x(k+1) = Ax(k) + Bu(k) + w(k) \tag{1a}$$

$$y^i(k) = C_i x(k) + v^i(k) + a^i(k); \quad k \in \mathbb{Z}_{\geq 0} \tag{1b}$$

where $x(k) \in \mathbb{R}^{n_x}$ is the state, $u(k) \in \mathbb{R}^{n_u}$ is a known input that is bounded, and $y^i(k) \in \mathbb{R}^{m_i}$ is the measured output of ith sensor with $i \in \mathbb{Z}_{[1,p]}$ and p the total number of sensors. The vector $w(k) \in \mathcal{W}$ represents the process noise, which is bounded and assumed to be contained in the zonotope $\mathcal{W} = \langle c_w, G_w\rangle$, and the vector $v^i(k) \in \mathcal{V}_i$ represents the measurement noise of ith sensor, which is also bounded and assumed to be contained in the zonotope $\mathcal{V}_i = \langle c_{v^i}, G_{v^i}\rangle$ for every sensor i. Finally, $a^i(k) \in \mathbb{R}^{m_i}$ represents the attack signal injected by the attacker to corrupt the measurement of ith sensor, and it can be arbitrary and unbounded.

Assumption 1.

(i) Upper bound on the number of attacked sensors: The attacker can attack up to $q < p$ number of sensors, where q is known a priori. However, the exact number and set of sensors which have been attacked are unknown.
(ii) Observability from each sensor: For every $i \in \mathbb{Z}_{[1,p]}$, (A, C_i) is an observable pair.
(iii) Knowledge of the initial set: The initial state $x(0)$ is contained in a zonotope $\mathcal{X}_0 = \langle c_0, G_0\rangle$.
(iv) Bounded input bounded state stability: There exists $M > 0$ such that $\|x(k)\|_\infty \leq M$, for any $w_{[0,k]} \in \mathcal{W} \subset \mathbb{R}^{n_x}$, and bounded $u_{[0,k]} \in \mathbb{R}^{n_u}$ with $k \in \mathbb{Z}_{\geq 1}$.

Assumption 1 (i) is fundamental in this paper because it ensures that, at every time $k \in \mathbb{Z}_{\geq 0}$, there exists a set of uncompromised sensors $\mathsf{S}_k \subset \mathbb{Z}_{[1,p]}$ with $|\mathsf{S}_k| = p - q$ such that $a^i(k) = 0_{m_i}$ for every $i \in \mathsf{S}_k$. This, along with Assumption 1(ii), allows us to ensure that the true state is included inside the intersection of the estimated sets of uncompromised sensors. In addition, the assumption entails that the attacker, even though omniscient about the system dynamics and noise bounds, has limited resources at hand. We remark that this assumption is certainly not restrictive, because it neither restricts the set of attacked sensors to be static with respect to time nor requires that q is less than half the number of sensors p. In contrast, at any time instant, the attacker can inject arbitrary attack signals to any subset of sensors with cardinality less than or equal to q, where q is only required to be strictly less than p.

Assumption 1 (ii) is required to enable decentralized set-based operations for resilient estimation without violating the robustness guarantees. Moreover, because Assumption 1(i) allows the attacker to attack up to $p-1$ sensors, it is necessary that the observability is guaranteed from any sensor.

Assumption 1 (iii) can be easily satisfied from the operating conditions of the system, and it is not restrictive because the size of $\mathcal{X}_0$ is not required to be small.

Finally, Assumption 1(iv) demarcates the class of systems considered in this paper and assumes bounded input bounded state (BIBS) stability, which is equivalent to saying that state matrix A is Schur stable (i.e., $\rho(A) < 1$). While this may appear to be restrictive in comparison to other resilient state estimation schemes for

[1] https://github.com/aalanwar/Secure-Set-Based-Estimation.

LTI systems where the A matrix does not need to be Schur stable (discrete-time systems), or Hurwitz (continuous-time systems), we argue that the class of BIBS stable systems is not restrictive as it is a property all control systems strive to achieve via feedback.

Under the standing assumptions stated above, we formulate the problem statement as follows.

Problem Statement

Given the model $A, C_1, \dots, C_p$, noise zonotopes $\mathcal{W}$ and $\mathcal{V}_1, \dots, \mathcal{V}_p$, output measurements $y^1(k), \dots, y^p(k)$, and the maximum number q of sensors that can be attacked at any time k, we aim to estimate a set $\hat{\mathcal{X}}_k$ guaranteeing the inclusion $x(k) \in \hat{\mathcal{X}}_k$ for every $k \in \mathbb{Z}_{\geq 0}$, where $x(k)$ is the true state of system (1).

This paper achieves the aforementioned resilient set-based state estimation problem via zonotopic filtering, which we will develop in the forthcoming sections.

4. Resilient zonotopic filtering

In this section, we propose our main algorithm for resilient set-based state estimation, which is summarized below.

4.1. Reachable set and time update step

The reachable set $\mathcal{R}_k$ at time $k \in \mathbb{Z}_{\geq 0}$ is the set of states to which the system may evolve given the input $u(k-1)$, and a guarantee that the previous state is contained in $\mathcal{R}_{k-1} \ni x(k-1)$ and the process noise in $\mathcal{W} \ni w(k-1)$, i.e.,

$$\mathcal{R}_k = A\mathcal{R}_{k-1} \oplus Bu(k-1) \oplus \mathcal{W}. \tag{2}$$

Because of the open-loop computation of the reachable set, i.e., it is not corrected using the sensor measurements, it turns out to be quite conservative for larger values of time k. Nonetheless, the Eq. (2) is particularly important in the time update step, also known as the prediction step, of our proposed filtering algorithm. The *time update* is given by

$$\hat{\mathcal{X}}_{k|k-1} = A\hat{\mathcal{X}}_{k-1|k-1} \oplus Bu(k-1) \oplus \mathcal{W} \tag{3}$$

where $\mathcal{R}_k$ is replaced by the time update $\hat{\mathcal{X}}_{k|k-1}$ and the previous reachable set $\mathcal{R}_{k-1}$ by the previous *measurement update* $\hat{\mathcal{X}}_{k-1|k-1}$. The measurement update, also known as the correction step, is described in Section 4.3. Notice that the attacker cannot directly influence the time update (3), but it can influence (3) indirectly through the measurement update $\hat{\mathcal{X}}_{k-1|k-1}$. Thus, it is important to carefully devise the measurement update, which we do in Section 4.3, for achieving resilience against sensor attacks.

4.2. State space region consistent with the measurements

Before presenting the measurement update, we estimate a subset of state space that is consistent with the sensor measurements. Given the output Eq. (1b), output matrix C_i, and the measurement noise bound $\mathcal{V}_i$, a method to find a subset of state space consistent with the sensor i's measurement $y^i(k)$ is provided in the following lemma, which is inspired by Alanwar et al. [3], 4]. To this end, we employ the singular value decomposition (SVD) of the output matrix of the ith sensor

$$C_i = \begin{bmatrix} P_1^i & P_2^i \end{bmatrix} \begin{bmatrix} \Sigma_{r_i} & 0_{r_i \times (n_x - r_i)} \\ 0_{(m_i - r_i)\times r_i} & 0_{(m_i-r_i)\times(n_x - r_i)} \end{bmatrix} \begin{bmatrix} V_1^{i\top} \\ V_2^{i\top} \end{bmatrix}$$

where $\operatorname{rank}(C_i) = r_i \leq m_i$, $P_1^i \in \mathbb{R}^{m_i \times r_i}$, $P_2^i \in \mathbb{R}^{m_i \times (m_i - r_i)}$, $V_1^i \in \mathbb{R}^{n_x \times r_i}$, $V_2^i \in \mathbb{R}^{n_x \times (n_x - r_i)}$, and $\Sigma_{r_i} \in \mathbb{R}^{r_i \times r_i}$ is a positive definite diagonal matrix. Then, the pseudo-inverse of C_i is given by $C_i^\dagger = V_1^i \Sigma_{r_i}^{-1} P_1^{i\top}$. If C_i is full row rank, i.e., $r_i = m_i$, then $C_i^\dagger = C_i^\top (C_i C_i^\top)$.

Lemma 1. *Let Assumption 1(ii) and (iv) hold. Then, for every $i \in \mathbb{Z}_{[1,p]}$, the state space region consistent with the measurement $y^i(k) = C_i x(k) + v^i(k) + a^i(k)$ is given by the zonotope*

$$\mathcal{Y}_k^i = \langle c_{y^i}(k), G_{y^i}(k) \rangle, \text{ where} \tag{4a}$$

$$\begin{cases} c_{y^i}(k) &= C_i^\dagger \left(y^i(k) - c_{v^i} \right) \\ G_{y^i}(k) &= [C_i^\dagger G_{v^i} \quad MV_2^i] \end{cases} \tag{4b}$$

where M is given by Assumption 1(iv).

Moreover, if $i \in \mathsf{S}_k$, where S_k is the set of uncompromised sensors at time k, then $x(k) \in \mathcal{Y}_k^i$, for every $k \in \mathbb{Z}_{\geq 0}$.

Proof. Since (A, C_i) is observable for every $i \in \mathsf{S}_k$, we have that the solution to

$$y^i(k) = C_i x(k) + v^i(k) = C_i x(k) + v^i(k) + x(k) - x(k)$$

given by

$$x(k) = C_i^\dagger (y^i(k) - v^i(k)) + (I_{n_x} - C_i^\dagger C_i) x(k)$$

is non-trivial, where $a^i(k) = 0_{m_i}$ since $i \in \mathsf{S}_k$. By Assumption 1(iv), $x(k) \in \langle 0, MI_{n_x} \rangle$, which implies that

$$x(k) \in C_i^\dagger (y^i(k) - \mathcal{V}) \oplus (I_{n_x} - C_i^\dagger C_i) \langle 0, MI_{n_x} \rangle$$

To show that the right hand side of the above inclusion equals $\mathcal{Y}_k^i$, note that

$$\operatorname{im}(V_2^i) = \operatorname{im}(I_{n_x} - C_i^\dagger C_i) = \ker(C_i).$$

Therefore, $(I_{n_x} - C_i^\dagger C_i) x(k) \in \operatorname{im}(V_2^i)$, implying

$$(I_{n_x} - C_i^\dagger C_i) \langle 0, MI_{n_x} \rangle = V_2^i \langle 0, MI_{n_x - r_i} \rangle.$$

Hence,

$$\begin{aligned} &C_i^\dagger (y^i(k) - \mathcal{V}) \oplus (I_{n_x} - C_i^\dagger C_i) \langle 0, MI_{n_x} \rangle \\ &\quad = C_i^\dagger (y^i(k) - \mathcal{V}) \oplus V_2^i \langle 0, MI_{n_x - r_i} \rangle \\ &\quad = C_i^\dagger (y^i(k) - \mathcal{V}) \oplus \langle 0, MV_2^i \rangle \\ &\quad = \mathcal{Y}_k^i \end{aligned}$$

which completes the proof. □

By Assumption 1(iv), the state space is given by the zonotope $\langle 0, MI_{n_x} \rangle$. Subject to this assumption, (4) in the above lemma computes a subset of $\langle 0, MI_{n_x} \rangle$ that is consistent with the sensor i's measurement. Thus, if sensor i is unattacked at time k, it is guaranteed that the true state $x(k)$ is inside the set $\mathcal{Y}_k^i$. However, the guarantee doesn't hold when i is under attack at time k. To verify if a subset of sensors is not attacked and can be trusted, it is necessary that the intersection of their consistent sets yields a non-empty set. This intersection will discard all the sensors whose measurements are corrupted by large attack signals. However, sensors that are injected by stealthy attack signals, i.e., signals within the noise bounds, remain undetected. Nonetheless, we can ensure that there is at least one subset of sensors with cardinality $p - q$ that is guaranteed to contain the true state $x(k)$.

Theorem 2. *Let Assumption 1 hold. Then, there exists an index set $\mathsf{J} \subset \mathbb{Z}_{[1,p]}$ with cardinality $|\mathsf{J}| = p - q$ such that $x(k) \in \mathcal{I}_k$, where $\mathcal{I}_k$ is a constrained zonotope given by*

$$\mathcal{I}_k = \bigcap_{i \in \mathsf{J}} \mathcal{Y}_k^i \tag{5}$$

with $\mathcal{Y}_k^i$ given in (4).

Proof. By Assumption 1(i), the number of uncompromised sensors $|\mathsf{S}_k| \geq p - q$, because the attacker can attack only up to q sensors. Thus, there exists $\mathsf{J} \subset \mathbb{Z}_{[1,p]}$ with cardinality $|\mathsf{J}| = p - q$ containing

only the uncompromised sensors, i.e., $\mathsf{J} \subseteq \mathsf{S}_k$. Since the inclusion $x(k) \in \mathcal{Y}_k^i$ for every $i \in \mathsf{S}_k$ is guaranteed by Lemma 1, and there exists J with cardinality $|\mathsf{J}| = p - q$ such that $\mathsf{J} \subseteq \mathsf{S}_k$, the inclusion $x(k) \in \mathcal{I}_k$ is guaranteed with $\mathcal{I}_k$ given in (5). □

We have shown that there exists a subset of sensors whose consistent sets yield a non-empty intersection, and the intersection contains the true state. However, in the presence of stealthy attacks, it is not possible to completely discard the attacked sensors. There could be multiple subsets of sensors whose consistent sets yield non-empty intersections, but only some of them may contain the true state.

4.3. Measurement update step

Measurement update $\hat{\mathcal{X}}_{k|k}$ corrects the conservative estimate of the model-based time update by incorporating new information from the sensor measurements (4). In other words, the measurement update step involves intersecting the time update set $\hat{\mathcal{X}}_{k|k-1}$ with the state space regions consistent with the sensor measurements.

4.3.1. Measurement update in the absence of attacks

First, consider the following result in the absence of the attacker.

Lemma 3. *Let Assumption 1 hold with the number of attacks $q = 0$ for every $k \in \mathbb{Z}_{\geq 0}$, i.e., $a^i \equiv 0_{m_i}$ for every $i \in \mathbb{Z}_{[1,p]}$. Then, for every $k \in \mathbb{Z}_{\geq 1}$, it holds that $x(k) \in \hat{\mathcal{X}}_{k|k}$, where $\hat{\mathcal{X}}_{k|k}$ is a constrained zonotope given by*

$$\hat{\mathcal{X}}_{k|k} = \hat{\mathcal{X}}_{k|k-1} \cap \left(\bigcap_{i=1}^{p} \mathcal{Y}_k^i \right) \tag{6}$$

and $\hat{\mathcal{X}}_{k|k-1}$ is given in (3) with $\hat{\mathcal{X}}_{1|0} = A\mathcal{X}_0 \oplus Bu(0) \oplus \mathcal{W}$.

Proof. Since $x(0) \in \mathcal{X}_0$ by Assumption 1(iii), we have that $x(1) \in \hat{\mathcal{X}}_{1|0}$. Also, by Lemma 1, $x(1) \in \mathcal{Y}_1^i$ for every $i \in \mathbb{Z}_{[1,p]}$. Therefore, we have $x(1) \in \hat{\mathcal{X}}_{1|1}$. This, in turn, implies that $x(2) \in \hat{\mathcal{X}}_{2|1}$. By applying Lemma 1 again, we have that $x(2) \in \hat{\mathcal{X}}_{2|2}$. Thus, by induction, for every $k \in \mathbb{Z}_{\geq 1}$, $x(k-1) \in \hat{\mathcal{X}}_{k-1|k-1}$ implies $x(k) \in \hat{\mathcal{X}}_{k|k-1}$, which guarantees $x(k) \in \hat{\mathcal{X}}_{k|k}$ by Lemma 1. □

The Eq. (6) is the usual measurement update in the absence of attacker, which is central to zonotopic filtering [12]. However, this measurement update may yield an empty estimated set $\hat{\mathcal{X}}_{k|k} = \emptyset$ even when only one sensor is under attack. In this case, the attacker has to only ensure that the attack signal is large enough so that the consistent sets yield an empty intersection. Therefore, when considering that a subset of sensors might be attacked, a more sophisticated way of performing measurement update is developed next.

4.3.2. Measurement update in the presence of attacks

To obtain the measurement update $\hat{\mathcal{X}}_{k|k}$ in the presence of attacker, we propose to intersect the time update $\hat{\mathcal{X}}_{k|k-1}$ with the state space regions consistent with the measurements of all subsets of sensors with cardinality $p - q$. That is, for every index set $\mathsf{J}_h \subset \mathbb{Z}_{[1,p]}$ with cardinality $|\mathsf{J}_h| = p - q$, compute the intersection of consistent sets

$$\mathcal{I}_k^h = \bigcap_{j \in \mathsf{J}_h} \mathcal{Y}_k^j \tag{7}$$

where $h = 1, \ldots, \eta$ with

$$\eta = \binom{p}{p-q} = \frac{p!}{q!(p-q)!}. \tag{8}$$

Then, the measurement update $\hat{\mathcal{X}}_{k|k}$ is obtained as

$$\hat{\mathcal{X}}_{k|k} = \hat{\mathcal{X}}_{k|k-1} \cap \{\mathcal{I}_k^h\}_{h \in \mathbb{Z}_{[1,\eta]}} \tag{9}$$

where we note that $\hat{\mathcal{X}}_{k|k}$ is a collection of multiple constrained zonotopes.

Theorem 4. *Let Assumption 1 hold. Then, given the measurement update $\hat{\mathcal{X}}_{k|k}$ from (9), the inclusion $x(k) \in \hat{\mathcal{X}}_{k|k}$ is guaranteed for every $k \in \mathbb{Z}_{\geq 1}$.*

Proof idea. The inclusion can be guaranteed through the same arguments as in the proof of Lemma 3 but using Theorem 2 instead of Lemma 1. □

Although the inclusion of the true state is guaranteed by the above theorem, it is important to remark that the number of sets in the measurement update (9) may increase with respect to time under stealthy attacks. We address this issue in Section 4.5 by proposing several techniques that facilitate computational efficiency of the algorithm.

It is worth mentioning that the proposed algorithm is resilient because the attacker cannot deteriorate the estimation accuracy over time. If a subset J_h contains a sensor which is injected by a large attack signal, it will be automatically discarded because of an empty intersection $\mathcal{I}_k^h$ in (7). Therefore, in order to yield a non-empty intersection, the attacker can only inject small attack signals whose magnitude is within the measurement noise bounds $\mathcal{V}_i$, which does not deteriorate the estimation accuracy.

4.4. Bound on the estimation error

Since Theorem 4 guarantees that the true state $x(k)$ of system (1) lies in at least one of the zonotopes in the measurement update $\hat{\mathcal{X}}_{k|k}$ at each $k \in \mathbb{Z}_{\geq 0}$, it must also lie in a zonotope that overbounds $\hat{\mathcal{X}}_{k|k}$. That is, we overbound the collection of constrained zonotopes in $\hat{\mathcal{X}}_{k|k}$ by another constrained zonotope $\hat{\mathcal{Z}}_k = \langle \hat{c}_z(k), \hat{G}_z(k) \rangle$, which is obtained by solving

$$\min \operatorname{rad}(\hat{\mathcal{Z}}_k) \text{ subject to } \hat{\mathcal{X}}_{k|k} \subset \hat{\mathcal{Z}}_k. \tag{10}$$

Then, the estimation error can be bounded by

$$\|\hat{c}_z(k) - x(k)\| \leq \operatorname{rad}(\hat{\mathcal{Z}}_k).$$

It can be proven that the error computed above is upper bounded asymptotically because of the stable time-update step (Section 4.1). Moreover, in practice, this error bound is significantly smaller than the error bounds obtained by point-based resilient estimators [13,16].

4.5. Methods to reduce the complexity

The major computational challenge of Algorithm 1 that can be exploited by the attacker lies in the measurement update step (9) for computing $\hat{\mathcal{X}}_{k|k}$, which is a collection of zonotopes whose cardinality (i.e., the number of zonotopes) could grow over time. To reduce computational complexity resulting from the increasing cardinality of the measurement update, we propose several pruning methods. The first step is to remove the empty sets or subsets of other sets in the measurement update intersection (9). It is also possible to obtain a single overbounding zonotope of $\hat{\mathcal{X}}_{k|k}$ as in (10), and use it in the next time update step. However, a better trade-off between accuracy and complexity is to not overbound the whole collection, but only the intersecting zonotopes in the collection $\hat{\mathcal{X}}_{k|k}$. This may not make the cardinality of $\hat{\mathcal{X}}_{k|k}$ equal to one, but it reduces it significantly by allowing minimal loss of accuracy.

Another method is employ a point-based resilient estimator, if it exists, in parallel with the set-based resilient estimator. In

Algorithm 1 Resilient zonotope-based state estimation.

Require: System matrices A, B, and C_i, and noise zonotopes $\mathcal{W}$ and $\mathcal{V}_i$, for every $i \in \mathbb{Z}_{[1,p]}$; time sequence of sensor measurements $\{y^1(k), y^2(k), \ldots, y^p(k)\}_{k\in\mathbb{Z}_{\geq 0}}$.

1: Initialize: $\hat{\mathcal{X}}_{0|0} = \mathcal{X}_0$
2: **for** $k = 1, 2, 3, \ldots$ **do**
3: Time update: $\hat{\mathcal{X}}_{k|k-1} = A\hat{\mathcal{X}}_{k-1|k-1} \oplus Bu(k-1) \oplus \mathcal{W}$
4: Obtain $\mathcal{Y}_k^i = \langle c_{y^i}(k), G_{y^i}(k)\rangle$ using (4), for every sensor $i \in \mathbb{Z}_{[1,p]}$.
5: Obtain $\mathcal{I}_k^h$ using (7) for every index set $\mathsf{J}_h \subset \mathbb{Z}_{[1,p]}$ with cardinality $|\mathsf{J}_h| = p - q$ and $h \in \mathbb{Z}_{[1,\eta]}$.
6: Measurement update: $\hat{\mathcal{X}}_{k|k} = \hat{\mathcal{X}}_{k|k-1} \cap \{\mathcal{I}_k^h\}_{h\in\mathbb{Z}_{[1,\eta]}}$
7: **end for**

this case, we may consider only those candidates in the measurement update collection that lie within the intersection of $\hat{\mathcal{X}}_{k|k}$ and an error margin generated by a point-based resilient state estimator. However, the existing point-based resilient state estimators [11,13,15,16,21] require that the total number of sensors be strictly greater than twice the number of compromised sensors $q < p/2$ and the members of the attacked sensors also remain unchanged over time, which are tighter requirements than our standing Assumption 1(i). Moreover, the error margins obtained by point-based estimators are usually very conservative.

Additionally, one may also employ zonotope reduction methods (see Yang and Scott [23] and the references therein) to reduce the number of generators in the zonotopes, which is often increased by the Minkowski sum operation. However, this technique may result in larger radius of $\hat{\mathcal{Z}}_k$ in (10).

5. Case studies

In this section, we discuss three scenarios to evaluate the detection mechanisms and complexity under the proposed resilient zonotope-based state estimation algorithm.

5.1. Detection under time-invariant attacks

A notable relaxation of the set-based state estimation scheme in this paper over other resilient schemes is Assumption 1(i), which allows the attacker to compromise a different set of sensors over time. However, in the case where the set of attacked sensors is time-invariant, we can detect the set of compromised sensors by identifying the $\mathcal{Y}_k^i$ in (4) which do not intersect with each other or the time update set $\hat{\mathcal{X}}_{k|k}$. To be precise, under time-invariant sensor attacks, a subset of compromised sensors can be detected over time by building the following index set

$$\mathcal{D}_k \doteq \{i \in \mathbb{Z}_{[1,p]} : \mathcal{Y}_{k'}^i \cap \hat{\mathcal{X}}_{k'|k'-1} = \emptyset, \text{ for } k' \leq k\} \tag{11}$$

where the cardinality $|\mathcal{D}_k|$ is a non-decreasing function of time k. Notice that the detector (11) may not detect stealthy attacks, where the magnitude of attack signals is within the measurement noise bound. Nonetheless, in certain cases, the attacker can be detected as illustrated in Fig. 1(c).

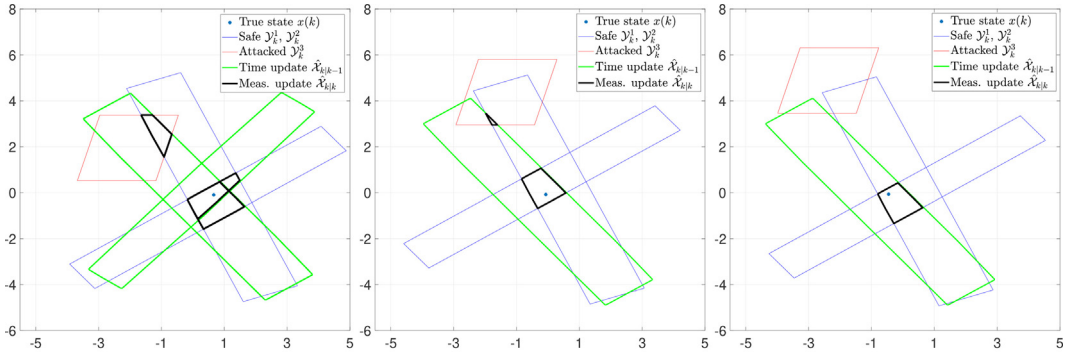
5.2. Naive attacks are discarded automatically

In the case where we have a naive attacker who injects large attack signals or random attack signals, the attacked sensors may be automatically discarded by our proposed Algorithm 1. As discussed already in Section 4.2, large attack signals are automatically discarded because they result in an empty intersection in (7). Random attack signals, even if within the noise bounds, can also be detected eventually if the attacker is not smart enough to discount for the changing orientation of the time update set and consistent sets of other sensors. Figure 1(c) illustrates such a scenario.

Sensor faults like denial of service (DoS) and intermittent transmissions come under naive attacks in our proposed framework, and they can be easily handled by Algorithm 1. Random attack signals injected by the attacker may result from their limited knowledge of the system or the noise bounds. It could also result from the fact that the attacker has limited resources at hand and cannot generate an optimal attack signal to ensure worst-case complexity at every time instant k. Under such assumptions, the attacked sensors can be discarded, which results in significant reduction of complexity of the measurement update step (9).

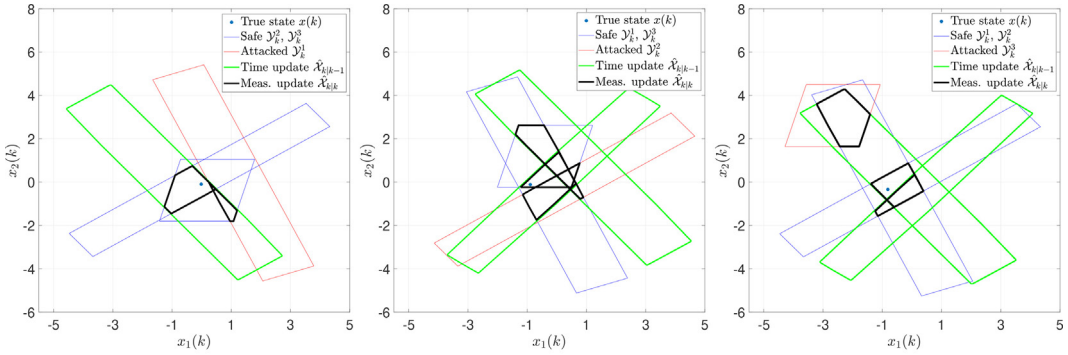
5.3. Stealthy attacks can increase complexity exponentially in the worst-case scenario

Stealthy attacks on sensors result in sets $\mathcal{Y}_k^i$, for $i \in \mathbb{Z}_{[1,p]} \setminus \mathsf{S}_k$, that may not yield any empty intersection in (7) in the worst-case scenario. This can increase the complexity of Algorithm 1, where the number of sets may increase exponentially with respect to time, which can overwhelm the available computation resources. To be precise, the number of zonotopes in the measurement update collection $\hat{\mathcal{X}}_{k|k}$ can be on the order of η^k in the worst-case scenario, where η is given in (8). Therefore, the methods discussed in Section 4.5 are very crucial to ensure computational feasibility of the proposed algorithm at the next time instant $k+1$. Since each complexity reduction method offers a trade-off between estimation accuracy and complexity, the best method is the one that offers maximum accuracy under the available computational resources.



(a) Time $k = 4$ (b) Time $k = 12$ (c) Time $k = 23$

Fig. 1. Snapshots of estimated sets at different times using Algorithm 1 under random attacks.



(a) Time $k = 3$, Sensor 1 attacked (b) Time $k = 4$, Sensor 2 attacked (c) Time $k = 5$, Sensor 3 attacked

Fig. 2. Snapshots of estimated sets using Algorithm 1 under time-varying attack, where different sensors are attacked at different times.

6. Evaluation

We evaluate our method by considering an input-driven variant of the rotating target described in Alanwar et al. [2]. The known input $u(k) \in \mathbb{R}$ is sampled uniformly at random from the set $\mathcal{U} = \langle 0, 10 \rangle$ at every time k. We have

$$A = \begin{bmatrix} 0.9455 & -0.2426 \\ 0.2486 & 0.9455 \end{bmatrix}, \quad B = \begin{bmatrix} 0.1 \\ 0 \end{bmatrix}. \tag{12}$$

With the number of sensors $p = 3$, we consider output matrices and respective measurement noise zonotopes as follows

$$C_1 = \begin{bmatrix} 1 & 0.4 \end{bmatrix}, \quad C_2 = \begin{bmatrix} 0.9 & -1.2 \end{bmatrix}, \quad C_3 = \begin{bmatrix} -0.8 & 0.2 \\ 0 & 0.7 \end{bmatrix}$$

$$\mathcal{V}_1 = \langle 0, 1 \rangle, \quad \mathcal{V}_2 = \langle 0, 1 \rangle, \quad \mathcal{V}_3 = \langle [0 \ \ 0]^\top, I_2 \rangle.$$

The process noise signal $w(k)$ are bounded by the zonotope $\mathcal{W} = \langle [0 \ \ 0]^\top, 0.02 I_2 \rangle$. The noise signals $v^i(k)$ and $w(k)$ are sampled uniformly at random from their respective zonotope sets using the function `randPoint(`$\mathcal{Z}$`)` in CORA [5]. The simulation was performed on the 11th Generation Intel(R) Core(TM) i7-1185G7 processor with 16.0 GB RAM, which took 0.273 s on average per iteration.

Fig. 1 presents three snapshots of the time-updated sets from the previous step (green), safe measurement consistent sets (blue), attacked measurement consistent set (red), and the final estimated measurement update sets (black). The time update sets (green) are computed using (2). Lemma 1 is used to compute the state space regions consistent with the measurements (blue) in which one of them is under attack (red). The measurement update sets (black) are computed according to (9). It is to be noted that the true state always remains inside the measurement update sets (Theorem 4).

Fig. 1 (a) and (b) show different scenarios in which the attacked set is intersecting with the intersection of the time update set and the safe sets. On the other hand, Fig. 1(c) shows a scenario in which the attacked set is not intersecting with the intersection of the safe and time update sets, which allows us to discard the attacked set and obtain a single measurement update set. Such a scenario may arise in non-intelligent stealthy attacks, where the attack signals are generated randomly at every time.

Fig. 2 shows a more powerful, time varying attack in which the attacker attacks a different sensor at different time steps. In Fig. 2(a), Sensor 1 is under attack and we have two estimated measurement update sets (black). Then, Sensor 2 is attacked in Fig. 2(b) in which the number of estimated measurement update sets is increasing due to having a small attack value. Finally, Sensor 3 is attacked in Fig. 2(c) with a larger attack value. Although the complexity increases in such attacks, it is worth noting that the true state $x(k)$ remains enclosed by the estimated measurement update sets at all time steps. Also, the estimation error remains bounded and the attacker cannot destroy the accuracy of the set-based state estimate.

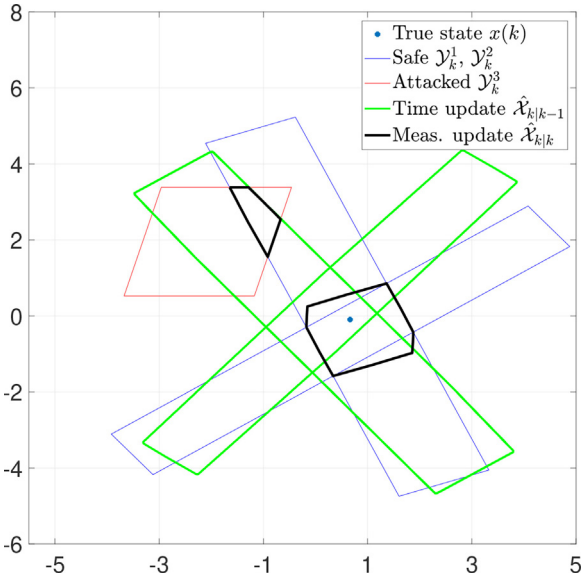


Fig. 3. Over-bounding two intersecting safe sets in Fig. 1(a) by a single set to reduce complexity.

Finally, the question of reducing the complexity by minimally compromising on the accuracy remains. To this end, Fig. 3 illustrates one of the complexity reduction methods discussed in Section 4.5, where we over-bound multiple intersecting sets by a single constrained zonotope. This significantly reduces the number of sets in the measurement update collection.

7. Conclusions and future outlook

We have presented a resilient zonotope-based state estimation scheme for LTI systems with multiple redundant sensors, i.e., the pair (A, C_i) is observable from every sensor i. We show that our scheme ensures that the true state lies within the estimated set. We acknowledge that the scheme suffers from the curse of dimensionality and we discuss complexity reduction methods. We discuss cases under which our proposed algorithm can be sharpened, by the design of a detection algorithm in the case where the set of attacked sensors remain constant and when the attacker performs naive attacks. On the other hand, stealthy attacks can also increase the complexity of the scheme, which underlines the importance of methods for reduction.

Future work will focus on relaxing the current observable via every sensor assumption to the case where the system is observable through a subset of sensors, which will increase the applicability of the scheme.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work is supported by the Swedish Research Council and the Knut and Alice Wallenberg Foundation, Sweden. It has also received funding from the European Union's Horizon Research and Innovation Programme under grant agreement no. 830927 and Marie Skłodowska-Curie grant agreement no. 101062523.

References

[1] A. Alanwar, H. Said, M. Althoff, Distributed secure state estimation using diffusion Kalman filters and reachability analysis, in: 2019 IEEE 58th Conference on Decision and Control (CDC), IEEE, 2019, pp. 4133–4139.
[2] A. Alanwar, J. J. Rath, H. Said, M. Althoff, Distributed set-based observers using diffusion strategy, (2020) arXiv:2003.10347.
[3] A. Alanwar, A. Berndt, K.H. Johansson, H. Sandberg, Data-driven set-based estimation using matrix zonotopes with set containment guarantees, in: 2022 European Control Conference (ECC), 2022a, pp. 875–881.
[4] A. Alanwar, M.U.B. Niazi, K.H. Johansson, Data-driven set-based estimation of polynomial systems with application to SIR epidemics, in: 2022 European Control Conference (ECC), 2022b, pp. 888–893.
[5] M. Althoff, An introduction to CORA 2015, in: Proceedings of the Workshop on Applied Verification for Continuous and Hybrid Systems, 2015.
[6] M. Althoff, J.J. Rath, Comparison of guaranteed state estimators for linear time-invariant systems, Automatica 130 (2021). Article no. 109662
[7] M. Althoff, G. Frehse, A. Girard, Set propagation techniques for reachability analysis, Annu. Rev. Control, Robot., Auton. Syst. 4 (2021) 369–395.
[8] D. Bertsekas, I. Rhodes, Recursive state estimation for a set-membership description of uncertainty, IEEE Trans. Autom. Control 16 (2) (1971) 117–128.
[9] J. Blesa, V. Puig, J. Saludes, Robust fault detection using polytope-based set-membership consistency test, IET Control Theory Appl. 6 (12) (2012) 1767–1777.
[10] P. Bouron, D. Meizel, P. Bonnifait, Set-membership non-linear observers with application to vehicle localisation, in: 2001 European Control Conference (ECC), 2001, pp. 1255–1260.
[11] M.S. Chong, H. Sandberg, J.P. Hespanha, A secure state estimation algorithm for nonlinear systems under sensor attacks, in: 2020 59th IEEE Conference on Decision and Control (CDC), 2020, pp. 5743–5748.
[12] A.A. de Paula, G.V. Raffo, B.O. Teixeira, Zonotopic filtering for uncertain nonlinear systems: fundamentals, implementation aspects, and extensions [applications of control], IEEE Control Syst. Mag. 42 (1) (2022) 19–51.
[13] X. He, X. Ren, H. Sandberg, K.H. Johansson, How to secure distributed filters under sensor attacks, IEEE Trans. Autom. Control 67 (6) (2021) 2843–2856.
[14] L. Jaulin, Robust set-membership state estimation: application to underwater robotics, Automatica 45 (2009) 202–206.
[15] J. Kim, C. Lee, H. Shim, Y. Eun, J.H. Seo, Detection of sensor attack and resilient state estimation for uniformly observable nonlinear systems having redundant sensors, IEEE Trans. Autom. Control 64 (3) (2018) 1162–1169.
[16] M. Pajic, I. Lee, G.J. Pappas, Attack-resilient state estimation for noisy dynamical systems, IEEE Trans. Control Netw. Syst. 4 (1) (2016) 82–92.
[17] B.S. Rego, S.G. Vrachimis, M.M. Polycarpou, G.V. Raffo, D.M. Raimondo, State estimation and leakage detection in water distribution networks using constrained zonotopes, IEEE Trans. Control Syst. Technol. 30 (5) (2021) 1920–1933.
[18] U. Shaked, Y. Theodor, $\mathcal{H}_\infty$-optimal estimation: a tutorial, in: [1992] Proceedings of the 31st IEEE Conference on Decision and Control, 1992, pp. 2278–2286.
[19] T. Shinohara, T. Namerikawa, Reach set-based attack resilient state estimation against omniscient adversaries, in: 2018 Annual American Control Conference (ACC), 2018a, pp. 5813–5818.
[20] T. Shinohara, T. Namerikawa, Reach set-based secure state estimation against sensor attacks with interval hull approximation, SICE J. Control, Meas., Syst. Integr. 11 (5) (2018b) 399–408.
[21] Y. Shoukry, P. Nuzzo, A. Puggelli, A.L. Sangiovanni-Vincentelli, S.A. Seshia, P. Tabuada, Secure state estimation for cyber-physical systems under sensor attacks: a satisfiability modulo theory approach, IEEE Trans. Autom. Control 62 (10) (2017) 4917–4932.
[22] D. Simon, Optimal State Estimation: Kalman, H_∞, and Nonlinear Approaches, John Wiley & Sons, 2006.
[23] X. Yang, J.K. Scott, A comparison of zonotope order reduction techniques, Automatica 95 (2018) 378–384.