Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

On the resilience of a class of Correntropy-based state estimators

Alexandre Kircher [*] **Laurent Bako** [*] **Éric Blanco** [*]
Mohamed Benallouch [**]

[*] *Université de Lyon, Laboratoire Ampère (Ecole Centrale Lyon, CNRS UMR 5005), F-69134 Lyon, France.*
[**] *Université de Lyon, ECAM Lyon, Lab ECAM, F-69321 Lyon, France.*

Abstract: This paper deals with the analysis of a class of offline state estimators for LTI discrete-time systems in the presence of an arbitrary measurement noise which can potentially take any value. The considered class of estimators is defined as the solution of an optimization problem involving a performance function which can be interpreted as a generalization of cost functions used in the Maximum Correntropy Criterion. The conclusion of the analysis is that if the system is observable enough, then the considered class of estimators is resilient, which means that the obtained estimation error is independent from the highest values of the measurement noise. In the case of systems with a bounded process noise, the considered class of estimators provides a bounded estimation error under the appropriate conditions despite not being designed for this scenario.

Keywords: Secure state estimation, Maximum Correntropy Criterion (MCC), optimal estimation, Cyber-physical systems.

1. INTRODUCTION

The problem considered in this paper is the estimation of the (hidden) state of a system despite the presence of an arbitrary measurement noise which can take any value. This kind of problem typically occurs when designing an estimator to cope with faulty sensors (Niedfeldt and Beard, 2014). Arbitrary noises are also a suitable way of modeling attacks of diverse natures, like replay attacks (Mo and Sinopoli, 2009), which pose some challenges in estimation scenarios where the data are collected via a communication network. This is case for example for Cyber-Physical Systems (Cardenas et al., 2008).

To address the estimation problem in the face of such challenging uncertainties, many approaches have been developed in the literature among which faulty measurements isolation (Mishra et al., 2017; Pasqualetti et al., 2013), compressive sampling-inspired methods (Fawzi et al., 2014; Pajic et al., 2017), or event-triggered resilient estimation (Shoukry and Tabuada, 2016; Liu et al., 2016).

In parallel to those methods, estimators based on the Maximum Correntropy Criterion (MCC) were developed to tackle the resilient state estimation problem. In order to design new filters more robust to non-Gaussian noises, the core idea is to use the correntropy as a similarity measure to design the cost function. However, there are several types of similarity measures and also several methods to derive the estimator. For instance, (Chen et al., 2017a; Liu et al., 2017) use Gaussian kernel functions as cost functions and derives a fixed-point algorithm in order to update the posterior estimation in a Kalman Filter. The scope of estimators inspired by the MCC then grew, including for example optimization-based estimators with cost functions under the form of a sum of exponential absolute value functions (Chen et al., 2017b), or online estimators which derive an MCC-based cost function under the form of a Kalman Filter with approximated weights based on the prior estimation at each step (Kulikova, 2017).

The MCC framework has already given very promising results, such as in machine learning with robust facial recognition (He et al., 2011) or in robust channel estimation for wireless communications (Ma et al., 2015). However, the question of the resilience of such algorithms is an open question which however gets more and more interest: in particular, (Bako, 2018) and (Chen et al., 2019) discussed this question under the scope of regression.

The goal of this paper is to give theoretical guarantees of resilience for a class of state estimators inspired by the kernel functions used in MCC-based approaches. The studied system model is given in the form of a LTI discrete-time state-space representation affected by an arbitrary measurement noise. There is indeed no assumption on its statistical properties, on its values or on the distribution of its largest values over time and among sensors. The main theoretical result of this paper states that under a certain sufficient condition on the observability of the system, the estimation error induced by those estimators is upper bounded by a value which does not depend on the largest values of the measurement noise. The obtained bound is therefore valuable in order to understand how the characteristics of the system impact the performances of the estimator as it is directly linked to the parameters of the system itself. It is however non-computable. We nonetheless provide simulation results for the discussed class of estimators which show good performances in the presence of arbitrary noise and even when a process noise is added in the state equation, for which it was not designed in the first place.

Outline. The rest of the paper is organized as follows. The system to be estimated will be introduced in Section 2 while the class of estimators will be defined in Section 3. Section 4 will be focused on deriving theoretical guarantees on the boundedness of the estimation error. Finally, Section 5 will provide simulation results to assess the performances of the class of estimators, in the case without process noise in subsection 5.1 and in the case with process noise in subsection 5.2; Section 6 provides some concluding remarks.

Notations. Throughout this paper, $\mathbb{R}_{\geq 0}$ (respectively $\mathbb{R}_{>0}$) designates the set of nonnegative (respectively positive) reals. We note $\mathbb{R}^a$ the set of (column) vectors with a real elements and for any vector z in $\mathbb{R}^a$, z_i with i in $\{1, ..., a\}$ is the i-th component of z. Moreover, $\mathbb{R}^{a\times b}$ is the set of real matrices with a rows and b columns. If $M \in \mathbb{R}^{a\times b}$, then $M^\top$ will designate the transposed matrix of M. $\|\cdot\|_2$ is the Euclidean norm, defined by $\|z\|_2 = \sqrt{z^\top z}$ for all z in $\mathbb{R}^a$. When applied to a matrix, $\|\cdot\|_2$ will designate the matrix norm induced by the Euclidean norm. For a finite set $\mathcal{S}$, the notation $|\mathcal{S}|$ will refer to the cardinality of $\mathcal{S}$, $\mathcal{S}^c$ the complementary set of $\mathcal{S}$ and $\mathcal{P}(\mathcal{S})$ will be the power set of $\mathcal{S}$, *i.e.* the collection of all subsets of $\mathcal{S}$.

2. STUDIED SYSTEM TO BE ESTIMATED

In this paper, we consider a Linear Time-Invariant discrete-time system

$$\Sigma : \begin{cases} x_{t+1} = Ax_t \\ y_t = Cx_t + f_t \end{cases} \tag{1}$$

with $A \in \mathbb{R}^{n\times n}$ the state transition matrix of the system and $C \in \mathbb{R}^{n_y \times n}$ the observation matrix of the system. $x_t \in \mathbb{R}^n$ is the state of the system at time $t \in \mathbb{Z}_+$; in particular $x_0 \in \mathbb{R}^n$ is called the initial state, while $y_t \in \mathbb{R}^{n_y}$ designates the output of the system. $\{f_t\} \subset \mathbb{R}^{n_y}$ is a noise sequence which can take potentially any value. A convenient way of describing it is as follows:

$$f_t = v_t + s_t, \tag{2}$$

where $\{v_t\}$ is a bounded white noise sequence whereas $\{s_t\}$ is a *sparse* noise sequence, meaning that most of its values are equal to zero but its non-zero values can be arbitrarily large. It is however important to understand that s_t is sparse with regards to both time and sensors indexes, which implies that $\{s_t\}$ can present several non-zero values in a row on a specific sensor. The bounded white noise v_t usually represents measurement noise in the output equation, while the sparse noise can represent many different things such as sensor failures, intermittent network outage or false data injection attacks.

Problem. In this paper, we consider the problem of estimating the state trajectory of system (1) on a fixed time horizon $\mathbb{T} = \{0, 1, ..., T-1\}$, which means estimating the matrix $X \triangleq (x_0\ x_1\ \cdots\ x_{T-1}) \in \mathbb{R}^{n\times T}$ with the measurement matrix $Y \triangleq (y_0\ y_1\ \cdots\ y_{T-1})$ containing T measurements and the model (1) of the system. The resulting estimated trajectory must be accurate even though the estimation process is undermined by the presence of f_t whose characteristics are already described above.

3. THE CLASS OF STATE ESTIMATORS

To address the problem stated above, we propose an optimization-based solution. To do so, we define the following cost function:

$$V_\Sigma(Y, z_0) = \sum_{(t,j)\in\mathbb{T}\times\mathbb{J}} e^{-\gamma\psi(y_{t,j} - \theta_{t,j}^\top z_0)} \tag{3}$$

with $\mathbb{J} = \{0, 1, ..., n_y\}$, $\gamma \in \mathbb{R}_{>0}$ a user-defined parameter and $z_0 \in \mathbb{R}^n$. For every $(t,j) \in \mathbb{T}\times\mathbb{J}$, $\theta_{t,j}^\top = c_j^\top A^t$ where $c_j^\top$ is the j-th row of the observation matrix C: every row vector $\theta_{t,j}^\top$ is a row of the observability matrix of system Σ over the time horizon $\mathbb{T}$, $\mathcal{O} = \left(C^\top\ (CA)^\top\ \cdots\ (CA^{T-1})^\top\right)^\top$. $\psi : \mathbb{R} \to \mathbb{R}$ is a real function which is assumed to verify the following properties:

(P1) **Positive definiteness:** $\psi(0) = 0$ and $\psi(a) > 0$ for all non-zero real a.
(P2) **Symmetry:** $\psi(-a) = \psi(a)$ for all $a \in \mathbb{R}$.
(P3) **Non-decreasingness:** for any a_1, a_2 in $\mathbb{R}$, $|a_1| < |a_2|$ implies $\psi(a_1) \leq \psi(a_2)$.
(P4) **Generalized Triangle Inequality (GTI):** there exists $\alpha \in \mathbb{R}_{>0}$ such that for all a_1, a_2 in $\mathbb{R}$,

$$\psi(a_1 - a_2) \geq \alpha\psi(a_1) - \psi(a_2). \tag{4}$$

Many functions verify this set of properties, such as every absolute value power function $a \mapsto |a|^p$ with $p \in \mathbb{N}$ (see (Bako, 2018)). It aims at generalizing the kernels used in Maximum Correntropy approaches (see (Chen et al., 2017a) for Gaussian kernel and (Chen et al., 2017b) for exponential absolute value kernels). The estimator can therefore be defined as the set-valued map $\Psi_\Sigma : \mathbb{R}^{n_y\times T} \to \mathcal{P}(\mathbb{R}^n)$ such that for any possible measurement matrix Y in $\mathbb{R}^{n_y\times T}$,

$$\Psi_\Sigma(Y) = \underset{z_0\in\mathbb{R}^n}{\arg\max}\, V_\Sigma(Y, z_0). \tag{5}$$

We can note that this rather defines *a class of estimators* which has as many members as there are ψ functions verifying (P1)–(P4).

To obtain an estimate of the whole trajectory, we then need to simulate system (1) with any estimated initial state $\hat{x}_0 \in \Psi_\Sigma(Y)$. From now on, the expression "estimated trajectories" will designate any matrix $\hat{X}(\hat{x}_0)$ such that

$$\hat{X}(\hat{x}_0) = \left(\hat{x}_0\ A\hat{x}_0\ \cdots\ A^{T-1}\hat{x}_0\right),\ \hat{x}_0 \in \Psi_\Sigma(Y). \tag{6}$$

With this estimation framework defined, our first goal is to provide a theoretical analysis of our estimator in order to exhibit its resilience properties.

4. THEORETICAL RESULTS

In this section, we will conduct an analysis which will result in an upper bound on the estimation error independent from the largest values of the arbitrary noise sequence $\{f_t\}$. This analysis is inspired by the one presented in (Bako, 2018) but in the state estimation framework. A discussion about the estimator performances on a system with process disturbances will also be provided.

4.1 Preliminaries

Before stating our main theoretical result, we will first introduce a few notations. For convenience, and without any loss of generality, we can assume that there is no (t,j) in $\mathbb{T}\times\mathbb{J}$ such that $\|\theta_{t,j}\|_2 = 0$[(1)]. We then define the following variable:

$$\sigma_\Sigma = \min_{(t,j)\in\mathbb{T}\times\mathbb{J}} \|\theta_{t,j}\|_2 > 0 \tag{7}$$

Moreover, given λ in $[0; 1]$ and $z_0 \in \mathbb{R}^n$, we also define the following set of indexes (t,j):

$$\mathcal{J}_\Sigma(z_0, \lambda) = \{(t,j) \in \mathbb{T}\times\mathbb{J} : |\theta_{t,j}^\top z_0| \geq \lambda\|\theta_{t,j}\|_2\|z_0\|_2\} \tag{8}$$

This set selects the rows in the observability matrix which are almost in the same direction as z_0: indeed, if (t,j) is in $\mathcal{J}_\Sigma(z_0, \lambda)$, then the vector $\theta_{t,j}$ is within the cone of direction z_0 and of half-top angle $\arccos(\lambda)$. Thus, $\lambda = 1$ is the case where $\theta_{t,j}$ and z_0 are colinear and $\lambda = 0$ is the case where $\theta_{t,j}$ can be in any direction. As a result, $|\mathcal{J}_\Sigma(z_0, \lambda)|$

[(1)] If $\mathcal{O}$ contains null rows, the following analysis can be conducted with a matrix collecting all the non-zero rows of $\mathcal{O}$.

can be considered as a local observability measure with regards to z_0 given that it represents how many $\theta_{t,j}$ are almost in the same direction as z_0 with a tolerance λ.

We also define a global observability parameter:

$$R_\Sigma(\lambda) = \inf_{z_0 \in \mathbb{R}^n} |\mathcal{J}_\Sigma(z_0, \lambda)|. \tag{9}$$

For any λ in $[0;1]$, $R_\Sigma(\lambda)$ will be between 0 and $n_y T$: in particular, if $\lambda = 0$, then $R_\Sigma(\lambda) = n_y T$. The higher $R_\Sigma(\lambda)$ will be for high values of λ, the more the system will be considered observable.

The following lemma provides more context about the link between $R_\Sigma(\lambda)$ and the observability of the system:

Lemma 1. Consider the system Σ defined in (1). This system is (i) observable, *i.e.* $\mathrm{rank}(\mathcal{O}) = n$, if and only if (ii) there exists $\lambda \in]0;1]$ such that $R_\Sigma(\lambda) \neq 0$.

Proof. $(i) \Rightarrow (ii)$: if the system is observable, let's assume that for all λ in $]0;1]$, $R_\Sigma(\lambda) = 0$. As the set $\{|\mathcal{J}_\Sigma(z_0, \lambda)|\}_{z_0 \in \mathbb{R}^n}$ is a subset of $\mathbb{N}$, its infimum is necessarily attained, which entails that there exists $a_\lambda \in \mathbb{R}^n$ such that $\mathcal{J}_\Sigma(a_\lambda, \lambda) = \emptyset$ for any $\lambda \in]0;1]$. For a fixed λ, $\mathcal{J}_\Sigma(a_\lambda, \lambda) = \emptyset$ implies that for all (t,j) in $\mathbb{T} \times \mathbb{J}$, $|\theta_{t,j}^\top a_\lambda| < \lambda \|\theta_{t,j}\|_2 \|a_\lambda\|_2$. By squaring the two sides of the inequalities and adding them for every (t,j), we obtain

$$a_\lambda^\top \mathcal{O}^\top \mathcal{O} a_\lambda < \lambda \|a_\lambda\|_2^2 \sum_{(t,j) \in \mathbb{T} \times \mathbb{J}} \|\theta_{t,j}\|_2^2 \tag{10}$$

From (Bernstein, 2009, Corollary 8.4.2), $a_\lambda^\top \mathcal{O}^\top \mathcal{O} a_\lambda \geq \sigma_{\min} a_\lambda^\top a_\lambda$ where $\sigma_{\min}$ designates the smallest eigenvalue of $\mathcal{O}^\top \mathcal{O}$. Since the system is observable, $\mathcal{O}^\top \mathcal{O}$ is positive definite which yields $\sigma_{\min} > 0$. As a result, (10) implies $\lambda > \lambda_e$ where $\lambda_e = \sqrt{\sigma_{\min} / \sum_{(t,j) \in \mathbb{T} \times \mathbb{J}} \|\theta_{t,j}\|_2^2} > 0$. This is a clear contradiction for any λ in $]0;\lambda_e] \cap]0;1]$, which proves the implication.

$(ii) \Rightarrow (i)$: by contraposition, if the system is not observable, then there exists $a \in \mathbb{R}^n \neq 0$ such that $\mathcal{O}a = 0$. Given that $\mathcal{O}a = \left(\theta_{0,1}^\top a \cdots \theta_{0,n_y}^\top a \; \theta_{1,1}^\top a \cdots \theta_{T-1,n_y}^\top a\right)^\top$, we have for every (t,j) in $\mathbb{T} \times \mathbb{J}$, $\theta_{t,j}^\top a = 0$. As a result, for any $\lambda \in]0;1]$, $\mathcal{J}_\Sigma(a, \lambda) = \emptyset$ since $|\theta_{t,j}^\top a| = 0$ cannot be greater than $\lambda \|\theta_{t,j}\|_2 \|a\|_2$: this leads to $R_\Sigma(\lambda) = 0$ for any $\lambda \in]0;1]$, which proves the implication.

This lemma shows that the observability is an equivalent condition to the existence of a $\lambda \in]0;1]$ such that $R_\Sigma(\lambda) \neq 0$: Σ will thus be supposed to be observable from now on.

Finally, we need to define a last notation. Given ε a positive real number, for any noise sequence $\{f_{t,j}\}$, it is possible to split the set $\mathbb{T} \times \mathbb{J}$ into two disjoint subsets,

$$\mathcal{I}_\varepsilon = \{(t,j) \in \mathbb{T} \times \mathbb{J} \, : \, |f_{t,j}| \leq \varepsilon\}, \tag{11}$$

which gathers the indexes of $f_{t,j}$ such that their absolute value is smaller than ε, and $\mathcal{I}_\varepsilon^c = \{(t,j) \in \mathbb{T} \times \mathbb{J} : |f_{t,j}| > \varepsilon\}$ which consists of the indexes of outliers in $\{f_{t,j}\}$ with regards to ε. ε therefore acts like a threshold we can choose and tune in order to conduct our analysis.

4.2 Main result

In the following theorem, we are going to express under which circumstances the norm of the estimation error on the initial state, *i.e.* $e_0 = \hat{x}_0 - x_0$, and consequently the estimation error over the whole trajectory $E = \hat{X} - X$, can be bounded by a value which does not depend on the largest values of the noise sequence $\{f_t\}$:

Theorem 2. Consider the state estimator (5) for system (1) under the assumption that its loss function ψ defined in (3) verifies properties (P1)–(P4). Let $\varepsilon > 0$. For any noise sequence $\{f_t\}$ and initial state x_0 in (1), generating a measurement matrix Y, such that

$$\frac{1}{1 + e^{-\gamma\psi(\varepsilon)}} R_\Sigma(\lambda) + e^{-\gamma\psi(\varepsilon)} |\mathcal{I}_\varepsilon| > n_y T, \tag{12}$$

is verified for some λ in $]0;1]$, the following holds true :

$$\forall \hat{x}_0 \in \Psi_\Sigma(Y), \; \psi(\lambda \sigma_\Sigma \|e_0\|_2) \leq \frac{1}{\gamma\alpha} \ln(1/\mu) \tag{13}$$

with $e_0 = \hat{x}_0 - x_0$ and

$$\begin{aligned}\mu = \frac{1 + e^{-\gamma\psi(\varepsilon)}}{|\mathcal{I}_\varepsilon| + R_\Sigma(\lambda) - n_y T} \Bigg[& \frac{1}{1 + e^{-\gamma\psi(\varepsilon)}} R_\Sigma(\lambda) \\ & + e^{-\gamma\psi(\varepsilon)} |\mathcal{I}_\varepsilon| - n_y T \Bigg]\end{aligned} \tag{14}$$

Moreover, if ψ is (strictly) increasing on $\mathbb{R}_{\geq 0}$, then it admits an invert function $\psi^{(-1)}$ and

$$N(\hat{X} - X) \leq \frac{M_\Sigma}{\lambda \sigma_\Sigma} \psi^{(-1)} \left(\frac{1}{\gamma\alpha} \ln(1/\mu) \right) \tag{15}$$

with $\hat{X}$ as defined in (6), $N(\hat{X} - X) = \max_{t \in \mathbb{T}} \|\hat{x}_t - x_t\|_2$ and M_Σ a constant depending on the system dynamics.

Proof. By definition (5) of the estimator Ψ_Σ, for any measurement matrix Y in $\mathbb{R}^{n_y \times T}$, we have

$$V_\Sigma(Y, x_0) \leq V_\Sigma(Y, \hat{x}_0) \tag{16}$$

for every $\hat{x}_0$ in $\Psi_\Sigma(Y)$. This yields

$$\sum_{(t,j) \in \mathbb{T} \times \mathbb{J}} e^{-\gamma\psi(f_{t,j})} \leq \sum_{(t,j) \in \mathbb{T} \times \mathbb{J}} e^{-\gamma\psi(y_{t,j} - \theta_{t,j}^\top \hat{x}_0)} \tag{17}$$

Left side sum in (17) will now be decomposed as $\mathbb{T} \times \mathbb{J} = \mathcal{I}_\varepsilon \cup \mathcal{I}_\varepsilon^c$ by definition (11) of $\mathcal{I}_\varepsilon$. If (t,j) is in $\mathcal{I}_\varepsilon$, then $|f_{t,j}| < \varepsilon$, which implies $\psi(f_{t,j}) \leq \psi(\varepsilon)$ because of (P3), and subsequently $e^{-\gamma\psi(f_{t,j})} \geq e^{-\gamma\psi(\varepsilon)}$. For any (t,j) in $\mathcal{I}_\varepsilon^c$, we also have $e^{-\gamma\psi(f_{t,j})} > 0$, which yields

$$|\mathcal{I}_\varepsilon| e^{-\psi(\varepsilon)} \leq \sum_{(t,j) \in \mathbb{T} \times \mathbb{J}} e^{-\gamma\psi(y_{t,j} - \theta_{t,j}^\top \hat{x}_0)} \tag{18}$$

Moreover, for all (t,j) in $\mathbb{T} \times \mathbb{J}$, we always have $\psi(y_{t,j} - \theta_{t,j}^\top \hat{x}_0) = \psi(\theta_{t,j}^\top x_0 + f_{t,j} - \theta_{t,j}^\top \hat{x}_0) = \psi(f_{t,j} - \theta_{t,j}^\top e_0)$ with $e_0 = \hat{x}_0 - x_0$. For (t,j) in $\mathcal{I}_\varepsilon$, we apply the GTI (4) to ψ, leading to $\psi(f_{t,j} - \theta_{t,j}^\top e_0) \geq \alpha\psi(\theta_{t,j}^\top e_0) - \psi(f_{t,j})$. For (t,j) in $\mathcal{I}_\varepsilon^c$, as ψ verifies (P1), we have $\psi(f_{t,j} - \theta_{t,j}^\top e_0) \geq 0$, so for any (t,j) in $\mathcal{I}_\varepsilon^c$, $e^{-\gamma\psi(f_{t,j} - \theta_{t,j}^\top e_0)} \leq 1$. As a result, we obtain

$$|\mathcal{I}_\varepsilon| e^{-\psi(\varepsilon)} \leq e^{\gamma\psi(f_{t,j})} \sum_{(t,j) \in \mathcal{I}_\varepsilon} e^{-\gamma\alpha\psi(\theta_{t,j}^\top e_0)} + |\mathcal{I}_\varepsilon^c|$$

$$\Leftrightarrow e^{-\gamma\psi(\varepsilon)} \left[|\mathcal{I}_\varepsilon| (1 + e^{-\psi(\varepsilon)}) - n_y T \right] \leq \sum_{(t,j) \in \mathcal{I}_\varepsilon} e^{-\gamma\alpha\psi(\theta_{t,j}^\top e_0)}$$

as $|\mathcal{I}_\varepsilon^c| = n_y T - |\mathcal{I}_\varepsilon|$ and $e^{\gamma\psi(f_{t,j})} < e^{\gamma\psi(\varepsilon)}$ for $(t,j) \in \mathcal{I}_\varepsilon$. Given λ between 0 and 1, we will now decompose $\mathcal{I}_\varepsilon$ depending on whether (t,j) belongs to $\mathcal{J}_\Sigma(e_0, \lambda)$ or not. If (t,j) is in $\mathcal{I}_\varepsilon \cap \mathcal{J}_\Sigma(e_0, \lambda)$, then $\psi(\theta_{t,j}^\top e_0) \geq \psi(\lambda \|\sigma_\Sigma\|_2 \|e_0\|_2)$. As $a \mapsto e^{-a}$ is decreasing, it yields

$$\begin{aligned} & e^{-\gamma\psi(\varepsilon)} \left[|\mathcal{I}_\varepsilon| (1 + e^{-\psi(\varepsilon)}) - n_y T \right] \\ & \leq |\mathcal{I}_\varepsilon \cap \mathcal{J}_\Sigma(e_0, \lambda)| \left[e^{-\gamma\alpha\psi(\lambda\sigma_\Sigma \|e_0\|_2)} - 1 \right] + |\mathcal{I}_\varepsilon| \end{aligned}$$

Moreover, we have

$$\begin{aligned}|\mathcal{I}_\varepsilon \cap \mathcal{J}_\Sigma(e_0,\lambda)| &= |\mathcal{I}_\varepsilon| + |\mathcal{J}_\Sigma(e_0,\lambda)| - |\mathcal{I}_\varepsilon \cup \mathcal{J}_\Sigma(e_0,\lambda)| \\ &\geq |\mathcal{I}_\varepsilon| + R_\Sigma(\lambda) - n_yT\end{aligned}$$

By choosing λ so that condition (12) is met, we obtain $|\mathcal{I}_\varepsilon| + R_\Sigma(\lambda) - n_yT > 0$ given that

$$|\mathcal{I}_\varepsilon| + R_\Sigma(\lambda) \geq \frac{1}{1+e^{-\gamma\psi(\varepsilon)}}R_\Sigma(\lambda) + e^{-\gamma\psi(\varepsilon)}|\mathcal{I}_\varepsilon| > n_yT. \tag{19}$$

Since $e^{-\gamma\psi(\lambda\sigma_\Sigma\|e_0\|_2)} - 1 \leq 0$, we obtain

$$\begin{aligned}e^{-\gamma\psi(\varepsilon)}&\left[|\mathcal{I}_\varepsilon|(1+e^{-\psi(\varepsilon)}) - n_yT\right] - |\mathcal{I}_\varepsilon| \\ &\leq (|\mathcal{I}_\varepsilon| + R_\Sigma(\lambda) - n_yT)\left[e^{-\gamma\alpha\psi(\lambda\sigma_\Sigma\|e_0\|_2)} - 1\right] \quad (20)\end{aligned}$$

Dividing both sides by $|\mathcal{I}_\varepsilon| + R_\Sigma(\lambda) - n_yT$ and simplifying the left hand side of the inequality then gives us

$$\mu \leq e^{-\gamma\alpha\psi(\lambda\sigma_\Sigma\|e_0\|_2)} \tag{21}$$

with μ as defined in (14). When condition (12) is met, μ is positive, so we can apply ln to both sides, yielding

$$\psi(\lambda\sigma_\Sigma\|e_0\|_2) \leq \frac{1}{\gamma\alpha}\ln(1/\mu) \tag{22}$$

If ψ is increasing on $\mathbb{R}_{\geq 0}$, then it is obviously invertible on that interval, which entails

$$\|e_0\|_2 \leq \frac{1}{\lambda\sigma_\Sigma}\psi^{(-1)}\left(\frac{1}{\gamma\alpha}\ln(1/\mu)\right) \tag{23}$$

Finally, with $\hat{X}$ as defined in (6), $\hat{x}_t = A^t\hat{x}_0$ and $x_t = A^tx_0$, so we have $\hat{x}_t - x_t = A^te_0$ for all $t \in \mathbb{T}$. Consequently, $N(\hat{X} - X) = \max_{t\in\mathbb{T}}\|A^te_0\|_2 \leq (\max_{t\in\mathbb{T}}\|A^t\|_2)\,\|e_0\|_2$. This eventually yields

$$N(\hat{X} - X) \leq \frac{\max_{t\in\mathbb{T}}\|A^t\|_2}{\lambda\sigma_\Sigma}\psi^{(-1)}\left(\frac{1}{\gamma\alpha}\ln(1/\mu)\right) \tag{24}$$

which is the desired result with $M_\Sigma = \max_{t\in\mathbb{T}}\|A^t\|_2$.

What this theorem states is that the estimator Ψ_Σ yields estimates which are bounded despite the presence of an arbitrary noise potentially taking any value. Unfortunately, condition (12) cannot be computed due to the fact that obtaining $R_\Sigma(\lambda)$ is a combinatorial problem and that having access to $|\mathcal{I}_\varepsilon|$ requires knowing the proportion of outliers in $\{f_{t,j}\}$ with regards to ε.

Nevertheless, it gives relevant information about what is important to ensure the resilience of the estimator. Indeed, the inequality is composed of three terms: on the left-hand side, there are two terms, one which depends on a quantitative observability of the system through $R_\Sigma(\lambda)$, one which depends on the number of reasonable measurements *with regards to* ε. The right-hand side there is composed of a constant term equal to the total number of measurements on the time horizon. To promote this condition, there are two important things:

- We need the observability criterion $R_\Sigma(\lambda)$ to be as large as possible
- The number of outliers must be somehow limited

The interpretation of condition (12) is that the estimation error is bounded if the system is observable enough, and the more it is observable, the more outliers the estimator is able to handle.

In addition, the bound itself in (15) gives information on what impacts the quality of the estimation: besides the conclusions obtained through the study of condition (12), we can see that the actual values of λ and ε play a role in the tightness of the bound. Given that it exists for any (λ, ε) which verify condition (12), the best bound obtained through Theorem 1 would be

$$N(E) \leq M_\Sigma \min_{\substack{(\lambda,\varepsilon)\\ \mu>0}}\left[\frac{1}{\lambda\sigma_\Sigma}\psi^{(-1)}\left(\frac{1}{\gamma\alpha}\ln(1/\mu)\right)\right] \tag{25}$$

4.3 Case of systems with dynamic bounded noise

We now consider a system Σ_w of the form

$$\Sigma_w : \begin{cases} x_{t+1} = Ax_t + w_t \\ y_t = Cx_t + f_t \end{cases} \tag{26}$$

where $\{w_t\}$ is a bounded noise sequence. Adding a noise component w_t in the state equation is a common way of modeling process disturbances which shifts the state from the dynamic induced by A. Even though our estimator is not designed for such systems, it can still be applied to Σ_w. Indeed, for any system Σ_w, there exists a system $\tilde{\Sigma}$ such that

$$\tilde{\Sigma} : \begin{cases} \tilde{x}_{t+1} = A\tilde{x}_t \\ y_t = C\tilde{x}_t + \tilde{f}_t \end{cases} \tag{27}$$

with $\tilde{x}_0 = x_0$, $\tilde{f}_t = C\tilde{w}_t + f_t$ and $\tilde{w}_t = \sum_{k=0}^{t-1} A^kw_k$. This new system verifies the structure defined in (1) and gives the exact same output as Σ_w. In addition, the gap between x_t and $\tilde{x}_t$ is equal to $x_t - \tilde{x}_t = A^tx_0 + \tilde{w}_t - A^tx_0 = \tilde{w}_t$. As a result, we can draw the following corollary from Theorem 2:

Corollary 3. Consider the state estimator (5) for system (1) under the assumption that its loss function ψ defined in (3) is invertible and verifies properties (P1)–(P4). Let $\varepsilon > 0$. For any noise sequence $\{f_t\}$ and initial state x_0 in (1), generating a measurement matrix Y, such that condition (12) is verified for some λ in $]0;1]$, the following holds true :

$$\begin{aligned}&\forall\hat{x}_0 \in \Psi_\Sigma(Y), \\ &N(\hat{X}(\hat{x}_0) - X) \leq \frac{M_\Sigma}{\lambda\sigma_\Sigma}\psi^{(-1)}\left(\frac{1}{\gamma\alpha}\ln(1/\mu)\right) + \max_{t\in\mathbb{T}}\|\tilde{w}_t\|_2\end{aligned} \tag{28}$$

with $\hat{X}(\hat{x}_0)$ as defined in (6), $\tilde{w}$ as defined in (27) and μ, N and M_Σ as defined in Theorem 2.

Proof. By applying Theorem 2 to system $\tilde{\Sigma}$ defined in (27), we obtain

$$\|e_0\|_2 \leq \frac{1}{\lambda\sigma_\Sigma}\psi^{(-1)}\left(\frac{1}{\gamma\alpha}\ln(1/\mu)\right) \tag{29}$$

In the case of the systems Σ_w, we have for all $t \in \mathbb{T}$, $e_t = \hat{x}_t - x_t = A^te_0 - \tilde{w}_t$, so by considering the norm N, we have

$$N(E) \leq M_\Sigma\|\hat{X} - X\|_2 + \max_{t\in\mathbb{T}}\|\tilde{w}_t\|_2 \tag{30}$$

which yields the desired result.

5. DISCUSSIONS ON THE IMPLEMENTATION OF THE ESTIMATOR

For our numerical tests, we considered a system (1) with the following parameters

$$A = \begin{pmatrix} 0.7 & 0.45 \\ -0.5 & 1 \end{pmatrix}, \quad C = (1\ \ 2)\,, x_0 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}. \tag{31}$$

In this system, f_t was decomposed as in (2): v_t is a uniformly distributed white noise, while s_t is a sparse vector. It is generated in two steps: (1) the indexes (t, j)

for which $s_{t,j} \neq 0$ are uniformly selected over $\mathbb{T} \times \mathbb{J}$, and then (2) for each (t, j) such that $s_{t,j} \neq 0$, the value of $s_{t,j}$ is set as the realization of a Gaussian process of variance 100 and mean 0.

The optimization problem which defines the estimator is differentiable, but non-convex, which makes its numerical implementation difficult. To assess its resilience properties, we designed an Iterative Reweighted Least Squares (IRLS) algorithm (see Appendix A) to try to solve the optimization problem when ψ is the square function, *i.e.* $\forall a \in \mathbb{R}$, $\psi(a) = a^2$. We also chose $\gamma = 1$. Note however that there is no theoretical evidence as to the convergence of this iterative process towards the true solution of the (nonconvex) optimization problem in (5).

To perform the following tests, each setting was realized 100 times and the obtained relative estimation errors $\|\hat{X} - X\|_2/\|X\|_2$ were averaged. The IRLS estimator, implementing the estimator Ψ_Σ discussed in this paper, and an Oracle Least Squares (OLS) estimator were implemented. The mention "Oracle" means it works on a version of y_t unaffected by sparse noise s_t, namely $y_{wt} = y_t - s_t$. The OLS was chosen as a reference for comparison given that it is the best estimator with regards to the covariance of the estimation error in presence of Gaussian noises (Geer, 2005, p. 1041) and should therefore provide good results with bounded noises. Additionally, to give a reference when the IRLS is presented in the case with process noise, we also implemented a ℓ_1-norm based estimator designed for the presence of disturbance noise and for which the resilience was discussed in our previous work (Kircher et al., 2020).

5.1 Sparsity test

First, we conduct a test in absence of w_t to assess how the proportion of outliers were affecting the performances of the estimator with a fixed signal-to-noise ratio (SNR) of 30dB for v_t. The obtained results are presented on Figure 1. As expected, the IRLS estimator maintains acceptable performances as the ratio of outliers in the sparse noise increases: its average relative estimation error is almost identical to the one of the OLS estimator until 70% of non-zero values in s_t.

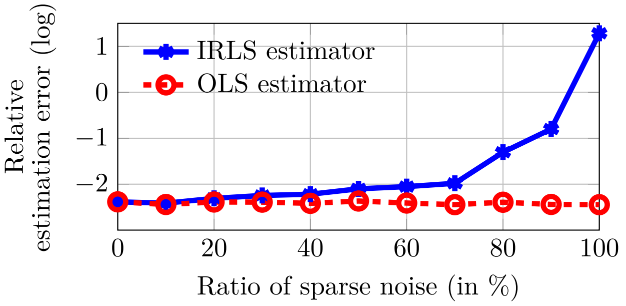

Fig. 1. Relative Estimation Error with different ratios of corrupted values for s_t and a SNR of 30dB for v_t

5.2 Performances of the estimator in presence of process noise

For the tests presented in this subsection we add a uniformly distributed white noise w_t in the state equation of the system. We then conduct two tests, one to see how the presence of a process noise w_t of 30dB degrades the performances of Ψ_Σ with increasing ratio of outliers in f_t (First test), and a second one to assess the impact of increasing SNR on the estimator (Second Test).

First test. The results obtained for the IRLS, ℓ_1-norm based and OLS estimators are gathered in Figure 2. The performance of the IRLS estimator are logically worse than in the previous subsection as a process noise w_t with a SNR of 40dB was introduced, and we see that the ℓ_1-norm based estimator is better on $]0; 60]$ given that its relative estimation error is closer to the one of the OLS estimator. However, as the analysis suggested, the performances of the IRLS estimator are still acceptable, and its estimation error seems bounded until 80% of outliers. It is even better than the ℓ_1-norm based estimator on the interval $[60; 90]$.

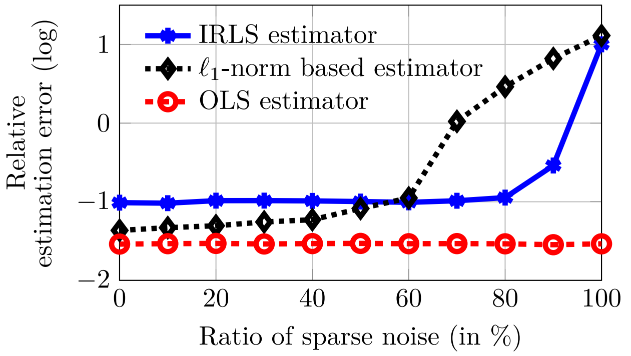

Fig. 2. Relative Estimation Error with different ratios of corrupted values for s_t and a SNR of 30dB for both w_t and v_t

Second Test. In this test, the ratio of outliers is fixed at 10%, while the SNR of both w_t and v_t simultaneously vary from 10dB to 100dB. Figure 3 displays the relative estimation error obtained with those different SNR for the three estimators. Despite the presence of 10% of outliers in s_t, the estimation error of both resilient estimators follow the same trend as the estimation error of the OLS one: it decreases as the SNR increases, which is logical given that noises are smaller compared to other quantities in the system equations. Nonetheless, once again, we observe that the IRLS estimator gives overall worse results than the ℓ_1-norm based one, which was to be expected considering it is outside of the scope of its design. Its performances are however still in the same range as the ℓ_1-norm based estimator.

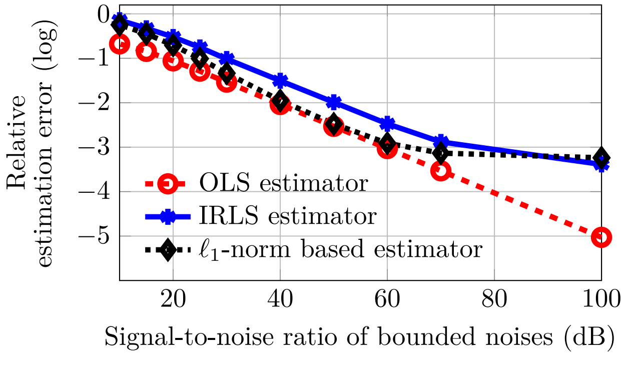

Fig. 3. Relative Estimation Error with varying SNR and 10% of corrupted values

6. CONCLUSION

In this paper, we have considered the problem of estimating the state of Linear Time-Invariant discrete-time systems in the face of uncertainties modeled as measurement noise in the system output equation. This noise sequence assumes values of possibly arbitrarily large amplitude which occur intermittently in time and across the available sensors. For this problem we have considered a class of state estimators through a cost function which generalizes the ones used in MCC-based approaches.

In particular, we have proven a resilience property for this class of state estimators so that the resulting estimation error is bounded by a value which is independent of the extreme values of the measurement noise provided that a certain condition linked to the observability of the system is met. Moreover, we proposed simulation results to assess the performance of this class, observing that it has good performances with and without the presence of a process noise. In future works, we would like to investigate the resilience property for an estimator defined by a cost function which would take into account the presence of process noise. Derivating theoretical guarantees for the designed IRLS algorithm would also be an interesting development.

Appendix A. ITERATIVE REWEIGHTED LEAST SQUARES (IRLS) ALGORITHM

This algorithm consists in solving a weigthed least mean square optimization problem

$$\min_{z_0 \in \mathbb{R}^n} V^{(i)}_{\Sigma,c}(Y, z_0) = \sum_{(t,j)\in\mathbb{T}\times\mathbb{J}} k^{(i)}_{t,j}(y_{t,j} - \theta^\top_{t,j} z_0)^2 \qquad \text{(A.1)}$$

with the weights $\{k^{(i)}_{t,j}\}$ being redefined after each step as

$$k^{(i+1)}_{t,j} = e^{-\gamma(y_{t,j}-\theta^\top_{t,j}x^{(i)}_0)^2} \qquad \text{(A.2)}$$

where $x^{(i)}_0$ refers to the solution of (A.1) obtain at step i. At each step, Problem (A.1) is solved through the CVX Solver in MATLAB (Grant and Boyd, 2018). When the relative difference between two consecutive estimated initial states is smaller than the user-defined threshold ε, the algorithm returns the state trajectory obtained from the last estimated state. The whole algorithm can be stated as follows:

Algorithm 1 Iterative Reweighted Least Squares (IRLS) Algorithm

```
Inputs: $\gamma$, $Y$, $\Sigma$, $\varepsilon$
Initialization:
$\forall (j,t) \in \mathbb{J}\times\mathbb{T}$, $k^{(1)}_{t,j} \leftarrow e^{-\gamma y^2_{t,j}}$
$\hat{x}^{(1)}_0 \leftarrow \arg\min_{z_0\in\mathbb{R}^n} V^{(1)}_{\Sigma,c}(Y, z_0)$
$\forall (t,j) \in \mathbb{T}\times\mathbb{J}$, $k^{(2)}_{t,j} \leftarrow e^{-\gamma(y_{t,j}-\theta^\top_{t,j}\hat{x}^{(1)}_0)^2}$
$\eta \leftarrow 10^8$
$i \leftarrow 1$
End of Initialization.
while $\eta > \varepsilon$ do
    $i \leftarrow i + 1$
    $\hat{x}^{(i)}_0 \leftarrow \arg\min_{z_0\in\mathbb{R}^n} V^{(i)}_{\Sigma,c}(Y, z_0)$
    $\forall (t,j) \in \mathbb{T}\times\mathbb{J}$, $k^{(i+1)}_{t,j} \leftarrow e^{-\gamma(y_{t,j}-\theta^\top_{t,j}\hat{x}^{(i)}_0)^2}$
    $\eta \leftarrow \frac{\|\hat{x}^{(i)}_0-\hat{x}^{(i-1)}_0\|}{\|\hat{x}^{(i-1)}_0\|}$
end while
return $\hat{X}(\hat{x}^{(i)}_0)$
```

REFERENCES

Bako, L. (2018). Robustness analysis of a maximum correntropy framework for linear regression. *Automatica*, 87, 218–225.

Bernstein, D.S. (2009). *Matrix Mathematics: Theory, Facts, and Formulas - Second Edition.* Princeton University Press. Publication Title: Matrix Mathematics.

Cardenas, A., Amin, S., and Sastry, S. (2008). Secure Control: Towards Survivable Cyber-Physical Systems. In *2008 The 28th International Conference on Distributed Computing Systems Workshops*, 495–500. IEEE.

Chen, B., Liu, X., Zhao, H., and Principe, J.C. (2017a). Maximum correntropy Kalman filter. *Automatica*, 76, 70–77.

Chen, B., Xing, L., Zhao, H., Du, S., and Príncipe, J.C. (2019). Effects of Outliers on the Maximum Correntropy Estimation: A Robustness Analysis. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 1–6.

Chen, Y., Ma, J., Zhang, P., Liu, F., and Mei, S. (2017b). Robust State Estimator Based on Maximum Exponential Absolute Value. *IEEE Transactions on Smart Grid*, 8(4), 1537–1544.

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.

Geer, S.A.v.d. (2005). Least squares estimation. In *Encyclopedia of Statistics in Behavioral Science*, 1041–1045. American Cancer Society.

Grant, M.C. and Boyd, S.P. (2018). The CVX Users' Guide.

He, R., Zheng, W.S., and Hu, B.G. (2011). Maximum Correntropy Criterion for Robust Face Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(8), 1561–1576.

Kircher, A., Bako, L., Blanco, E., and Benallouch, M. (2020). Analysis of resilience for a State Estimator for Linear Systems. *2020 American Control Conference.*

Kulikova, M.V. (2017). Square-root algorithms for maximum correntropy estimation of linear discrete-time systems in presence of non-Gaussian noise. *Systems & Control Letters*, 108, 8–15.

Liu, Q., Wang, Z., Liu, W., and Li, W. (2016). Event-triggered resilient filtering with missing measurements. In *2016 22nd International Conference on Automation and Computing (ICAC)*, 162–167.

Liu, X., Qu, H., Zhao, J., and Chen, B. (2017). State space maximum correntropy filter. *Signal Processing*, 130, 152–158.

Ma, W., Qu, H., Gui, G., Xu, L., Zhao, J., and Chen, B. (2015). Maximum correntropy criterion based sparse adaptive filtering algorithms for robust channel estimation under non-Gaussian environments. *Journal of the Franklin Institute*, 352(7), 2708–2727.

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 Sinopoli, B. (2009). Secure control against replay attacks. In *2009 47th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 911–918. IEEE.

Niedfeldt, P.C. and Beard, R.W. (2014). Robust estimation with faulty measurements using recursive-RANSAC. In *53rd IEEE Conference on Decision and Control*, 4160–4165. ISSN: 0191-2216.

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.

Pasqualetti, F., Dorfler, F., and Bullo, F. (2013). Attack Detection and Identification in Cyber-Physical Systems. *IEEE Transactions on Automatic Control*, 58(11), 2715–2729.

Shoukry, Y. and Tabuada, P. (2016). Event-Triggered State Observers for Sparse Sensor Noise/Attacks. *IEEE Transactions on Automatic Control*, 61(8), 2079–2091.