Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 58-15 (2024) 163–168

Lasso-based state estimation for cyber-physical systems under sensor attacks

V. Cerone [*] S. M. Fosson [*] D. Regruto [*] F. Ripa [*]

[*] *Department of Control and Computer Engineering,*
Politecnico di Torino, Italy (e-mail: sophie.fosson@polito.it).

Abstract. The development of algorithms for secure state estimation in vulnerable cyber-physical systems has been gaining attention in the last years. A consolidated assumption is that an adversary can tamper a relatively small number of sensors. In the literature, block-sparsity methods exploit this prior information to recover the attack locations and the state of the system. In this paper, we propose an alternative, Lasso-based approach and we analyse its effectiveness. In particular, we theoretically derive conditions that guarantee successful attack/state recovery, independently of established time sparsity patterns. Furthermore, we develop a sparse state observer, by starting from the iterative soft thresholding algorithm for Lasso, to perform online estimation. Through several numerical experiments, we compare the proposed methods to the state-of-the-art algorithms.

Keywords: Cyber-physical systems, sensor attacks, secure state estimation, sparse optimization, Lasso, Luenberger-like observers

1. INTRODUCTION

A cyber-physical system (CPS) is a collection of computing devices that interact with the physical world, through sensors and actuators, and with one another, through communication networks. Applications of the CPS paradigm include industrial control processes, smart power grids, wireless sensor networks, electric ground vehicles and cooperative driving technologies.

The distributed nature of CPSs is ambivalent in terms of security: on the one hand, it is more resilient to faults with respect to centralized systems; on the other hand, it is exposed to adversaries, either in terms of physical access to sensors or cyber access to data transmission networks. Examples range from non-invasive spoofing physical attacks, as illustrated by Shoukry et al. (2013), to cyber attacks to SCADA systems as done by the Stuxnet worm; see Langner (2011). Since CPSs act on the physical world, cyber-physical attacks layer may yield serious consequences to physical processes and to human beings, which makes the security problem very critical.

In the last decade, a substantial work has focused on the security of CPSs. A relevant research line considers the problem of secure state estimation (SSE) for CPSs in the presence of sensor attacks, that inject false data to manipulate the measurements. As a matter of fact, a malicious injection perturbs the data as a noise. However, we expect that an adversary conceives an unpredictable intrusion, that is, we have no information on its dynamics, boundedness and probabilistic description. The unique realistic assumption on sensor attacks is sparsity: only a relatively small number of sensors is accessible, due to, e.g., large dimensionality, physical deployment and sensor heterogeneity of CPSs.

In this paper, we consider CPSs described by discrete-time (DT) linear time-invariant (LTI) dynamical systems, with sparse sensor attacks. The identification of the attack support, i.e., the subset of tampered sensors, is a combinatorial problem, which does not scale well for large dimensional systems. However, by leveraging the sparsity assumption, one can exploit ℓ_1-based sparsity-promoting decoders to recast the problem into constrained convex optimization; see, e.g., Fawzi et al. (2014); Pajic et al. (2017). Since these approaches are still computationally intense, Shoukry and Tabuada (2016) introduce a faster event-triggered projected gradient (ETPG) approach, whose structure is prone to recursive SSE. The provided sufficient conditions for the convergence of ETPG are quite restrictive. Shoukry et al. (2017) address this issue by a satisfiability modulo theory approach, called Imhotep-SMT, which returns the exact solution in short time for problems of small/medium dimensions. However, Imhotep-SMT is combinatorial, thus critical for large-scale problems.

In this paper, we propose a different approach to SSE of CPSs under sparse sensor attacks, based on Lasso; see Tibshirani (1996). Specifically, we define a Lasso formulation and we analyse its effectiveness.

The proposed Lasso-based model builds on ℓ_1 relaxation, but, differently from Fawzi et al. (2014); Pajic et al. (2017), it gives rise to an unconstrained optimization problem, which can be solved through low-complex recursive algorithms. By elaborating on this point, the second contribution of the paper is a recursive SSE method exploiting

[⋆] This paper is part of the project NODES which has received funding from the MUR – M4C2 1.5 of PNRR with grant agreement no. ECS00000036.

10.1016/j.ifacol.2024.08.522

new data as soon as they become available, with the final aim of performing online SSE. More precisely, we design a sparsity-promoting Luenberger-like observer by starting from the iterative soft thresholding algorithm for Lasso.

Finally, we propose numerical experiments that show the effectiveness of the proposed methods with respect to the state-of-the-art approaches, in terms of estimation accuracy and execution time.

We organize the paper as follows. In Sec. 2, we state the problem and we illustrate the background. In Sec. 3, we introduce the proposed Lasso approach and we theoretically analyse it. In Sec. 4, we extend the method to recursive and online SSE, by developing a state observer. Finally, we devote Sec. 5 to numerical experiments and we draw some conclusions in Sec. 6.

2. PROBLEM STATEMENT

By following Fawzi et al. (2014); Shoukry and Tabuada (2016), we consider CPSs that can be modeled as DT LTI dynamical systems

$$\begin{aligned} x(k+1) &= Ax(k) \\ y(k) &= Cx(k) + a(k) \end{aligned} \qquad (1)$$

where $x(k) \in \mathbb{R}^n$ is the state, $y(k) \in \mathbb{R}^p$ is the measurement vector, $a(k) \in \mathbb{R}^p$ is the attack vector, $A \in \mathbb{R}^{n,n}$ and $C \in \mathbb{R}^{p,n}$. We assume that each sensor i takes a measurement $y_i(k)$; if $a_i(k) \neq 0$, sensor i is under attack. We assume that $a(k)$ is sparse, i.e., few sensors are under attack at each k. Since the presence of a known input does not impact on the formulation of the problem, see Shoukry and Tabuada (2016), for simplicity of notation we consider a zero-input model. Moreover, as in Fawzi et al. (2014); Shoukry and Tabuada (2016), we consider a noise-free model for our theoretical analysis, while we envisage measurement noise in some numerical experiments.

The SSE problem is as follows.

Problem 1. For some $\tau \leq n$ and $k \geq \tau - 1$ given A, C and $y = (y(k-\tau+1)^\top, \dots, y(k)^\top)^\top \in \mathbb{R}^{p\tau}$, estimate the τ-delayed state $x(k-\tau+1)$ in the presence of sparse sensor attacks.

Let us denote $\widetilde{a} = (a(k-\tau+1)^\top, \dots, a(k)^\top)^\top \in \mathbb{R}^{p\tau}$ and $\widetilde{x} = x(k-\tau+1) \in \mathbb{R}^n$, while $I \in \{0,1\}^{p\tau,p\tau}$ is the identity matrix. From (1), we have

$$y = (\mathcal{O}\ I)\begin{pmatrix} \widetilde{x} \\ \widetilde{a} \end{pmatrix}, \text{ where } \mathcal{O} = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{\tau-1} \end{pmatrix} \in \mathbb{R}^{p\tau,n} \qquad (2)$$

If $\tau = n$, $\mathcal{O}$ is the observability matrix of the attack-free system. We assume that the attack-free system is observable, i.e., $\operatorname{rank}(\mathcal{O}) = n$. In principle, we can estimate $(\widetilde{x}, \widetilde{a})$ by solving

$$y = (\mathcal{O}\ I)\begin{pmatrix} x \\ a \end{pmatrix} \qquad (3)$$

in the variables $x \in \mathbb{R}^n$, $a \in \mathbb{R}^{p\tau}$. Nevertheless, since $(\mathcal{O}\ I) \in \mathbb{R}^{p\tau,n+p\tau}$, the system is inherently underdetermined. However, by taking into account the sparsity of $\widetilde{a}$, we can exploit the compressed sensing theory to find a sparse solution to (3); see, e.g., Foucart and Rauhut (2013). Some works assume that the attack support is time-invariant, with cardinality $s \ll p$, i.e., $\widetilde{a}$ is block-sparse. In particular, Fawzi et al. (2014) exploit an ℓ_1/ℓ_r norm approach for block-sparse signal recovery, while Shoukry and Tabuada (2016) develop a block-based hard thresholding algorithm, that alternates gradient descent and event-triggered projection onto $\mathbb{R}^n \times \mathbb{S}_s$, where $\mathbb{S}_s \subset \mathbb{R}^{p\tau}$ is the set of block s-sparse vectors.

We remark that equation (3) has a unique solution in $\mathbb{R}^n \times \mathbb{S}_s$ if and only if the CPS defined in (1) is $2s$-sparse observable, that is, by removing any subset of $2s$ sensors, the attack-free system remains observable; see, e.g., (Shoukry and Tabuada, 2016, Theorem 3.2). In other terms, $2s$-sparse observability is equivalent to the injectivity of the map $f : \mathbb{R}^n \times \mathbb{S}_s \mapsto \mathbb{R}^{p\tau}$ defined as $f(x,a) = \mathcal{O}x + a$. This condition is necessary condition for secure state estimation, independently from the estimation algorithm.

If on the one hand the prior information on the block-sparsity pattern of $\widetilde{a}$ can improve the state estimation, on the other hand it ties the solution to constant attack supports and it originates more complex recovery algorithms. For these motivations, in this work we propose a sparse optimization approach that neglects the possible block-sparsity.

Beyond the seminal contributions briefly described in this section, we remark that substantial recent work addresses specific aspects of SSE of CPSs of the kind (1). For example, Mao et al. (2022) and Lu and Yang (2023) develop decomposition techniques, Zhao et al. (2023) propose a data-driven approach; Lei et al. (2023); Mao and Tabuada (2023) focus on distributed algorithms for SSE.

3. LASSO APPROACH

The proposed Lasso formulation for problem (3) is

$$(x^\star, a^\star) = \operatorname*{argmin}_{x \in \mathbb{R}^n, a \in \mathbb{R}^{p\tau}} \frac{1}{2}\|y - \mathcal{O}x - a\|_2^2 + \lambda\|a\|_1 \qquad (4)$$

where $\lambda > 0$. Problem (4) is a partial Lasso because only a part of the vector to estimate is sparse, i.e., we apply the ℓ_1 regularization only on a.

We notice that constrained counterparts of Lasso are the bases for the ℓ_1-based, block-sparsity decoders proposed by Fawzi et al. (2014), for the noise-free case, and by Pajic et al. (2017), in the presence of bounded measurement noise. The Lasso formulation in (4) may envisage the presence of measurement noise as well; moreover, we can resort to several effective algorithms for unconstrained optimization to solve it.

In this work, we consider the iterative soft thresholding algorithm (ISTA) proposed by Daubechies et al. (2004), which is a proximal gradient algorithm. ISTA iteration consists of a gradient step and a componentwise soft thresholding operation, defined by $\mathrm{S}_{\nu\lambda}[w] = w - \nu\lambda \operatorname{sign}(w)$ if $|w| \geq \nu\lambda$, and 0 otherwise; $\nu > 0$ is the gradient step size. Accelerated versions of ISTA are available, see, e.g., Beck and Teboulle (2009) and Cerone et al. (2023). Its simple structure allows us to build a state observer upon it, as illustrated in Sec. 4.

3.1 Analysis of the irrepresentable condition

An interesting feature of Lasso is that there is a tight condition, denoted as “irrepresentable”, that guarantees the recovery of the correct support in the noise-free case, see Fuchs (2004). Extensions to the noisy measurements are possible as well, see, e.g., Fuchs (2005).

In a nutshell, given a classic Lasso $\frac{1}{2}\|y-Qz\|_2^2+\lambda\|z\|_1$, the irrepresentable condition states that the columns of the sensing matrix Q on the true support must be “sufficiently orthogonal” to the columns outside the support. The irrepresentable condition cannot be priorly checked, because it involves the knowledge of the support; however, it provides interesting insights on the necessary features of Q to recover the support through Lasso; see, e.g., Zhao and Yu (2006); Hastie et al. (2015); Cerone et al. (2020).

In the considered SSE problem, the sensing matrix $(\mathcal{O}\ I)$ has a peculiar structure, with an identity matrix in its right part. In this section, we perform an irrepresentable condition analysis that takes into account this structure. We focus on the attack support recovery because its correct estimation allows us to recover the state from the safe sensors.

Let $\mathcal{S}$ be the support of $\widetilde{a}$, and $|\mathcal{S}| = h \ll p\tau$. If the attack support is time-invariant, $h = s\tau$. In the rest of the paper, we use the following notation: $I_{\mathcal{S}} \in \{0,1\}^{p\tau,h}$ is the submatrix of I with columns indexed in $\mathcal{S}$. $\mathcal{O}_{\mathcal{S}} \in \mathbb{R}^{h,n}$ and $\mathcal{O}_{\bar{\mathcal{S}}} \in \mathbb{R}^{p\tau-h,n}$ are the submatrices of $\mathcal{O}$ with *rows* in $\mathcal{S}$ and in $\bar{\mathcal{S}}$, respectively. $\mathcal{O}_{\bar{\mathcal{S}}}^{\dagger\top} = \mathcal{O}_{\bar{\mathcal{S}}}[\mathcal{O}_{\bar{\mathcal{S}}}^\top\mathcal{O}_{\bar{\mathcal{S}}}]^{-1} \in \mathbb{R}^{p\tau-h,n}$ is the right pseudo-inverse of $\mathcal{O}_{\bar{\mathcal{S}}}^\top$. Finally, we denote by $\|\cdot\|_\infty$ the ℓ_∞ matrix norm, defined as $\|M\|_\infty = \max_i \sum_j |M_{i,j}|$ for any matrix M. If M is a row vector, the ℓ_∞ matrix norm corresponds to the ℓ_1 vector norm.

The following result holds.

Theorem 1. Let us assume that $(\mathcal{O}\ I_{\mathcal{S}}) \in \mathbb{R}^{p\tau,n+h}$ is full rank. Lasso is successful, i.e., by solving it we identify the attack support, if and only if

$$\left\|\mathcal{O}_{\bar{\mathcal{S}}}^{\dagger\top}\mathcal{O}_{\mathcal{S}}^\top \mathrm{sign}(\widetilde{a}_{\mathcal{S}})\right\|_\infty < 1 \tag{5}$$

provided that $\lambda > 0$ is sufficiently small. As a consequence, Lasso is successful if

$$\left\|\mathcal{O}_{\bar{\mathcal{S}}}^{\dagger\top}\mathcal{O}_{\mathcal{S}}^\top\right\|_\infty < 1. \tag{6}$$

Proof. $(x^\star, a^\star)$ is a global minimum of (4) if and only if it fulfills the zero-subgradient condition: there exists $\zeta \in \partial\|a^\star\|_1$ such that

$$\begin{pmatrix}\mathcal{O}^\top\\ I\end{pmatrix}(\mathcal{O}x^\star + a^\star - y) + \lambda\begin{pmatrix}0\\ \zeta\end{pmatrix} = 0 \tag{7}$$

where $\partial\|a^\star\|_1 \subset \mathbb{R}^{p\tau}$ is the subdifferential of $\|a^\star\|_1$: for each $i \in \mathcal{S}$, $\zeta_i = \mathrm{sign}(a_i^\star)$, while $\|\zeta_{\bar{\mathcal{S}}}\|_\infty \le 1$; see Fuchs (2004) for details. In particular, if $\|\zeta_{\bar{\mathcal{S}}}\|_\infty < 1$, then $(x^\star, a^\star)$ is a strict minimum.

We construct the candidate $(x^\star, a^\star) \in \mathbb{R}^{n+p\tau}$ as follows: we set $a_{\bar{\mathcal{S}}}^\star = 0$, while

$$(x^\star, a_{\mathcal{S}}^\star) = \underset{x\in\mathbb{R}^n, a_{\mathcal{S}}\in\mathbb{R}^h}{\mathrm{argmin}}\ \frac{1}{2}\|y - \mathcal{O}x - I_{\mathcal{S}}a_{\mathcal{S}}\|_2^2 + \lambda\|a_{\mathcal{S}}\|_1 \tag{8}$$

which is a Lasso restricted on the non-zero components of the true vector. The so-built $(x^\star, a^\star)$ is a solution of Lasso (4) if it satisfies (7); let us verify under which conditions this occurs.

By distinguishing the zero-subgradient equations on $(x^\star, a_{\mathcal{S}}^\star)$ and on $a_{\bar{\mathcal{S}}}^\star$ and by recalling that $y = \mathcal{O}\widetilde{x} + I_{\mathcal{S}}\widetilde{a}_{\mathcal{S}}$, we have

$$(\mathcal{O}\ I_{\mathcal{S}})^\top(\mathcal{O}\ I_{\mathcal{S}})\begin{pmatrix}x^\star - \widetilde{x}\\ a_{\mathcal{S}}^\star - \widetilde{a}_{\mathcal{S}}\end{pmatrix} + \lambda\begin{pmatrix}0\\ \mathrm{sign}(a_{\mathcal{S}}^\star)\end{pmatrix} = 0 \tag{9}$$

$$I_{\bar{\mathcal{S}}}^\top(\mathcal{O}\ I_{\mathcal{S}})\begin{pmatrix}x^\star - \widetilde{x}\\ a_{\mathcal{S}}^\star - \widetilde{a}_{\mathcal{S}}\end{pmatrix} + \lambda\zeta_{\bar{\mathcal{S}}} = 0. \tag{10}$$

We notice that

$$(\mathcal{O}\ I_{\mathcal{S}})^\top(\mathcal{O}\ I_{\mathcal{S}}) = \begin{pmatrix}\mathcal{O}^\top\mathcal{O} & \mathcal{O}_{\mathcal{S}}^\top\\ \mathcal{O}_{\mathcal{S}} & I_h\end{pmatrix} \tag{11}$$

where $I_h \in \{0,1\}^{h,h}$ is the identity matrix of dimension h. Then, from (9) and (11), we compute

$$\begin{pmatrix}x^\star - \widetilde{x}\\ a_{\mathcal{S}}^\star - \widetilde{a}_{\mathcal{S}}\end{pmatrix} = -\lambda\begin{pmatrix}\mathcal{O}^\top\mathcal{O} & \mathcal{O}_{\mathcal{S}}^\top\\ \mathcal{O}_{\mathcal{S}} & I_h\end{pmatrix}^{-1}\begin{pmatrix}0\\ \mathrm{sign}(a_{\mathcal{S}}^\star)\end{pmatrix}. \tag{12}$$

The inverse matrix exists because the rows of $(\mathcal{O}\ I_{\mathcal{S}})^\top \in \mathbb{R}^{n+h,p\tau}$, $n+h < p\tau$, are linearly independent by assumption. Moreover, from (12), the difference $a_{\mathcal{S}}^\star - \widetilde{a}_{\mathcal{S}}$ is proportional to λ; therefore, if λ is sufficiently small, $\mathrm{sign}(a_{\mathcal{S}}^\star) = \mathrm{sign}(\widetilde{a}_{\mathcal{S}})$. We remark that, in the noise-free case, we can design λ as arbitrarily small without loss of generality.

By replacing (12) in (10), we obtain;

$$\zeta_{\bar{\mathcal{S}}} = I_{\bar{\mathcal{S}}}^\top(\mathcal{O}\ I_{\mathcal{S}})\begin{pmatrix}\mathcal{O}^\top\mathcal{O} & \mathcal{O}_{\mathcal{S}}^\top\\ \mathcal{O}_{\mathcal{S}} & I_h\end{pmatrix}^{-1}\begin{pmatrix}0\\ \mathrm{sign}(\widetilde{a}_{\mathcal{S}})\end{pmatrix}. \tag{13}$$

Now, $(x^\star, a^\star)$ is the unique minimum of Lasso if $\|\zeta_{\bar{\mathcal{S}}}\|_\infty < 1$. Therefore, we study the inequality

$$\left\|I_{\bar{\mathcal{S}}}^\top(\mathcal{O}\ I_{\mathcal{S}})\begin{pmatrix}\mathcal{O}^\top\mathcal{O} & \mathcal{O}_{\mathcal{S}}^\top\\ \mathcal{O}_{\mathcal{S}} & I_h\end{pmatrix}^{-1}\begin{pmatrix}0\\ \mathrm{sign}(\widetilde{a}_{\mathcal{S}})\end{pmatrix}\right\|_\infty < 1. \tag{14}$$

Since $I_{\bar{\mathcal{S}}}^\top I_{\mathcal{S}} = 0$,

$$I_{\bar{\mathcal{S}}}^\top(\mathcal{O}\ I_{\mathcal{S}}) = (\mathcal{O}_{\bar{\mathcal{S}}}\ 0). \tag{15}$$

From Schur's complement arguments,

$$\begin{pmatrix}\mathcal{O}^\top\mathcal{O} & \mathcal{O}_{\mathcal{S}}^\top\\ \mathcal{O}_{\mathcal{S}} & I_h\end{pmatrix}^{-1} = \begin{pmatrix}\Omega_1 & \Omega_2\\ \Omega_3 & \Omega_4\end{pmatrix} \tag{16}$$

where

$$\Omega_2 = -[\mathcal{O}^\top\mathcal{O} - \mathcal{O}_{\mathcal{S}}^\top\mathcal{O}_{\mathcal{S}}]^{-1}\mathcal{O}_{\mathcal{S}} = -[\mathcal{O}_{\bar{\mathcal{S}}}^\top\mathcal{O}_{\bar{\mathcal{S}}}]^{-1}\mathcal{O}_{\mathcal{S}}^\top. \tag{17}$$

By applying (15),(16) and (17) to (14),

$$(\mathcal{O}_{\bar{\mathcal{S}}}\ 0)\begin{pmatrix}\Omega_1 & \Omega_2\\ \Omega_3 & \Omega_4\end{pmatrix}\begin{pmatrix}0\\ \mathrm{sign}(\widetilde{a}_{\mathcal{S}})\end{pmatrix} = -\mathcal{O}_{\bar{\mathcal{S}}}^{\dagger\top}\mathcal{O}_{\mathcal{S}}^\top\mathrm{sign}(\widetilde{a}_{\mathcal{S}}). \tag{18}$$

In conclusion, condition (14) is equivalent to (5).

Since $\left\|\mathcal{O}_{\bar{\mathcal{S}}}^{\dagger\top}\mathcal{O}_{\mathcal{S}}^\top\mathrm{sign}(\widetilde{a}_{\mathcal{S}})\right\|_\infty \le \left\|\mathcal{O}_{\bar{\mathcal{S}}}^{\dagger\top}\mathcal{O}_{\mathcal{S}}^\top\right\|_\infty$ then (6) is sufficient for a successful Lasso. □

Since $\rho := \left\|\mathcal{O}_{\bar{\mathcal{S}}}^{\dagger\top}\mathcal{O}_{\mathcal{S}}^\top\right\|_\infty = \max_j\left\|\mathcal{O}_{\mathcal{S}}\left(\mathcal{O}_{\bar{\mathcal{S}}}^\dagger\right)_j\right\|_1$, a qualitative interpretation of (6) is as follows: the term ρ must be small, i.e., the rows of $\mathcal{O}_{\mathcal{S}}$ must be “sufficiently orthogonal” to the columns of $\mathcal{O}_{\bar{\mathcal{S}}}^\dagger$. This implies some observations. Since $\widetilde{x} = \mathcal{O}_{\bar{\mathcal{S}}}^\dagger y_{\bar{\mathcal{S}}}$, then

$$\mathcal{O}_{\mathcal{S}}\widetilde{x} = \mathcal{O}_{\mathcal{S}}\mathcal{O}_{\bar{\mathcal{S}}}^{\dagger} y_{\bar{\mathcal{S}}}. \tag{19}$$

In particular, $\mathcal{O}_{\mathcal{S}}\widetilde{x} = 0$ would be the ideal case to identify the attacks, which would be directly observable from $y_{\mathcal{S}} = a_{\mathcal{S}}$. Nevertheless, this is not realistic, in particular because it depends on the specific initial state $\widetilde{x}$. However,

$$\left\|\mathcal{O}_{\mathcal{S}}\widetilde{x}\right\|_1 \leq \rho \left\|y_{\bar{\mathcal{S}}}\right\|_1 \tag{20}$$

that is, ρ controls the "orthogonality" between the rows of $\mathcal{O}$ indexed in $\mathcal{S}$ and $\widetilde{x}$: a small irrepresentable term ρ implies a small energy of $\mathcal{O}_{\mathcal{S}}\widetilde{x}$, which, in turn, implies that the attacks are more exposed to identification.

Differently from previous conditions considered in the literature, see for example Theorem 4.4 in Shoukry and Tabuada (2016), the irrepresentable condition well captures the ability of CPS to identify attacks. We illustrate this point with a simple illustrative example.

Example 1. Let us consider a simple static model with $A = I$, $n = 1$ and $\mathcal{O} = C = \alpha\,(1\ 1\ 1)^{\top}$ for any $\alpha \in \mathbb{R}$, i.e., we have 3 equivalent sensors. We assume that one of them is under attack. If we know that $h = 1$, the identification of the attack is trivial: the sensor that provides a different measurement is clearly under attack and we can recover the state from the other two sensors. The irrepresentable condition well captures this resilience; in fact, we have $\mathcal{O}_{\bar{\mathcal{S}}}^{\dagger\top} = \frac{1}{2\alpha}\,(1\ 1)^{\top}$ and $\mathcal{O}_{\mathcal{S}}^{\top} = \alpha$, thus $\left\|\mathcal{O}_{\bar{\mathcal{S}}}^{\dagger\top}\mathcal{O}_{\mathcal{S}}^{\top}\right\|_{\infty} = \frac{1}{2} < 1$ for any $\alpha \in \mathbb{R}$ and Lasso is successful.

In contrast, we notice that condition 2 of Theorem 4.4 in Shoukry and Tabuada (2016) that guarantees the convergence of ETPG is not satisfied. In fact, the maximum eigenvalue of $(\mathcal{O}\ I)^{\top}(\mathcal{O}\ I)$ is $q = 3\alpha^2 + 1$, while the minimum eigenvalue on all the possible subsets of 2 sensors is $r = \frac{3\alpha^2+1-\sqrt{9\alpha^4+2\alpha^2+1}}{2}$. Then, $r < \frac{4}{9}q$, which contradicts condition 2.

4. SPARSE SOFT OBSERVER FOR ONLINE SSE

In this section, we move towards recursive, online SSE. We consider Problem 1 in a dynamic perspective: we aim at estimating the current state, or a delayed version, using the last $p\tau$ measurements, and, at each k we include the new measurements and we discard the oldest ones. If $\tau = 1$, this an online (not delayed) SSE. This calls for fast recursive online algorithms.

Shoukry and Tabuada (2016) address this problem by developing a recursive version of ETPG, named ETPL. At time k, instead of running ETPG to convergence to estimate $x(k - \tau + 1)$, ETPL runs some steps of the algorithm, then it moves to step $k + 1$ and it takes new measurements, in the philosophy of a Luenberger observer. As an alternative, we develop an online version of ISTA, that we name sparse soft observer and that we summarize in Alg. 1. We use the following notation: $\mathbf{a}(k) = (a(k-\tau+1)^{\top}, \dots, a(k)^{\top})^{\top}$, $\mathbf{y}(k) = (y(k-\tau+1)^{\top}, \dots, y(k)^{\top})^{\top}$.

The basic idea is to run one ISTA step, as described in Sec. 3, at each time instant k. Then, we update the estimate of the state by leveraging the knowledge of A, as in Luenberger observer. Step 3. in Alg. 1 can be repeated more than one time to enhance the estimation, but we do not run the complete algorithm to converge to the Lasso solution. Moreover, in case of online estimation, i.e., $\tau = 1$, the current p measurements are expected to be insufficient to have a successful Lasso; therefore, the observer approach is necessary.

Algorithm 1 Sparse soft observer

```
Input: $\tau \leq n$, $\lambda > 0$, $\nu > 0$, $A$, $\mathcal{O}$, $\mathbf{y}(k)$
Output: $\begin{pmatrix}\hat{x}(k)\\ \hat{\mathbf{a}}(k)\end{pmatrix}$ = estimate of $\begin{pmatrix}x(k-\tau+1)\\ \mathbf{a}(k)\end{pmatrix}$
1: for all $k = \tau - 1, \tau, \dots$ do
2:     Measurements and estimated measurements update
           $\mathbf{y}(k) = \mathcal{O}x(k-\tau+1) + \mathbf{a}(k)$
           $\hat{\mathbf{y}}(k) = \mathcal{O}\hat{x}(k) + \hat{\mathbf{a}}(k)$        (21)
3:     ISTA step: gradient step + soft thresholding
           $\begin{pmatrix}\hat{x}^{+}\\ \hat{\mathbf{a}}^{+}\end{pmatrix} = \begin{pmatrix}\hat{x}(k)\\ \hat{\mathbf{a}}(k)\end{pmatrix} - \nu\left(\mathcal{O}\ I\right)^{\top}\left[\hat{\mathbf{y}}(k) - \mathbf{y}(k)\right]$        (22)
           $\hat{\mathbf{a}}(k+1) = \mathrm{S}_{\nu\lambda}\left[\hat{\mathbf{a}}^{+}\right]$        (23)
4:     State update
           $\hat{x}(k+1) = A\hat{x}^{+}$        (24)
5: end for
```

5. NUMERICAL RESULTS

In this section, we propose some numerical results to investigate the performance of the proposed Lasso approach and sparse soft observer, in terms of state estimation accuracy and execution time. We perform all the simulations in MATLAB R2023 on a processor i7 @ 1.80 GHz × 8, with 16 GB of RAM.

5.1 Lasso approach

We test Lasso, solved by FISTA, see Beck and Teboulle (2009), on random, synthetic CPSs and we compare it to ETPG by Shoukry and Tabuada (2016) and Imhotep-SMT by Shoukry et al. (2017); the code for this last one is taken at "Imhotep-SMT" (2015). We generate the elements A and C independently, according to a standard normal distribution; then we normalize A to guarantee stability. We assume that the attack support is time-invariant and cardinality s and we generate it uniformly at random. The initial state $x(0)$ has uniformly distributed components with magnitude in $[2, 3]$. The attacks have magnitude in $[4, 5]$, which is sufficiently large to sabotage the state estimation, but not enough large to produce clear, plainly detectable outliers in the measurements.

In the first experiment, we vary p, while $n = 20$ and $s = \frac{p}{5}$; in the second experiment, we vary s, while $n = 20$ and $p = 30$. We perform 50 runs for each experiment; we depict the averages results in Fig. 1 and in Fig. 2. We assess the accuracy in terms of state estimation error $\|\hat{x} - \widetilde{x}\|_2 / \|\widetilde{x}\|_2$. We consider either noise-free and noisy measurements, i.e., $y(k) = Cx(k) + a(k) + \eta(k)$, where $\eta(k) \in \mathbb{R}^p$ is a uniformly random, bounded noise with $\|\eta(k)\|_{\infty} \leq 10^{-4}$.

As we can see in Fig. 1 and in Fig. 2, in this experiment, Lasso outperforms ETPG both in accuracy and run time. Since we consider small/medium dimensions, Imhotep-SMT is the best approach to provide the exact solution in fast time, in the noise-free case; nevertheless, its accuracy

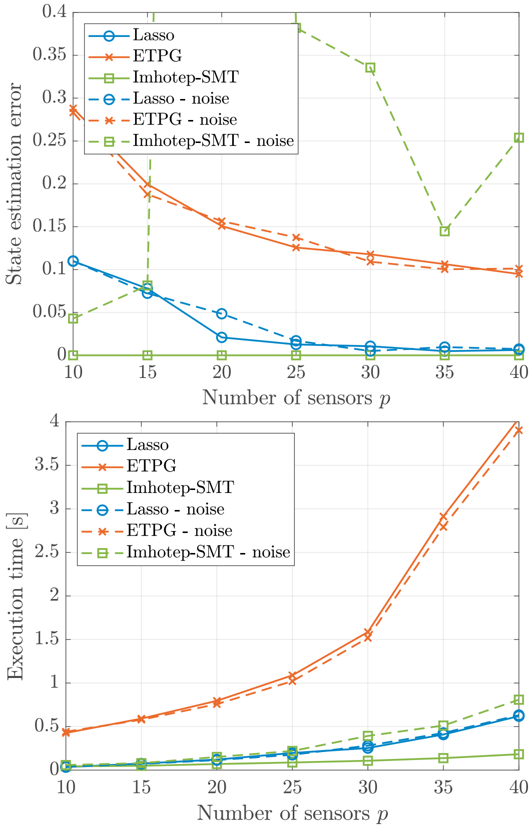

Figure 1. Lasso vs ETPG vs Imhotep-SMT, $n = 20$, $s = p/5$, noise-free and with noise bound 10^{-4}

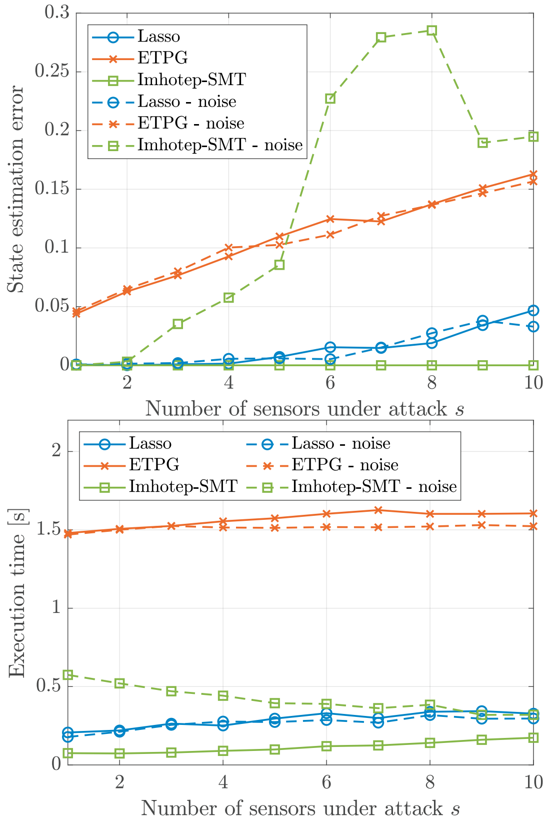

Figure 2. Lasso vs ETPG vs Imhotep-SMT, $n = 20$, $p = 30$, noise-free and with noise bound 10^{-4}

is critical in the presence of small measurement noise. In contrast, Lasso and ETPG are robust to small noise, with a very slight degradation in the estimation accuracy.

5.2 Sparse soft observer

In this section, we test the proposed sparse soft observer for recursive and online SSR. We perform 100 random runs and we show the average results, in terms of time evolution of the state estimation error $\|\hat{x} - \widetilde{x}\|_2/\|\widetilde{x}\|_2$ and support error, defined as $\sum_j |\mathbf{1}(\hat{\mathbf{a}}_j \neq 0) - \mathbf{1}(\widetilde{a}_j \neq 0)|$, where $\mathbf{1}(v) = 1$ if v is true and 0 otherwise.

We generate noise-free dynamical models as in the previous experiments, with $n = 10$, $p = 15$, $s = 3$, and we run each system for 300 time steps. We implement the proposed sparse soft observer and, for comparison, ETPL by Shoukry and Tabuada (2016). We show the results in Fig. 3 and in Fig. 4. In Fig. 3, we set $\tau = n$, that is, at each time step k we use the previous $p\tau$ measurements to estimate the delayed $x(k - \tau + 1)$. Then, at each k, we remove the oldest p measurements to introduce the new p ones. This is the case considered also by Shoukry and Tabuada (2016). Instead, in Fig. 4, we consider $\tau = 1$, that is, the algorithms use the current set of measurements $y(k)$ to provide an online estimate of $x(k)$. In the sparse soft observer, we repeat the step 3 of Alg. 1. for 5τ times. We can see that ETPL is more accurate for $\tau = n$, in particular it converges more quickly than the sparse soft observer to the correct attack support. However, the execution time at each k is $2 \cdot 10^{-3}$ seconds for ETPL and $4 \cdot 10^{-4}$ seconds for the sparse soft observer, so the latter is adaptable to

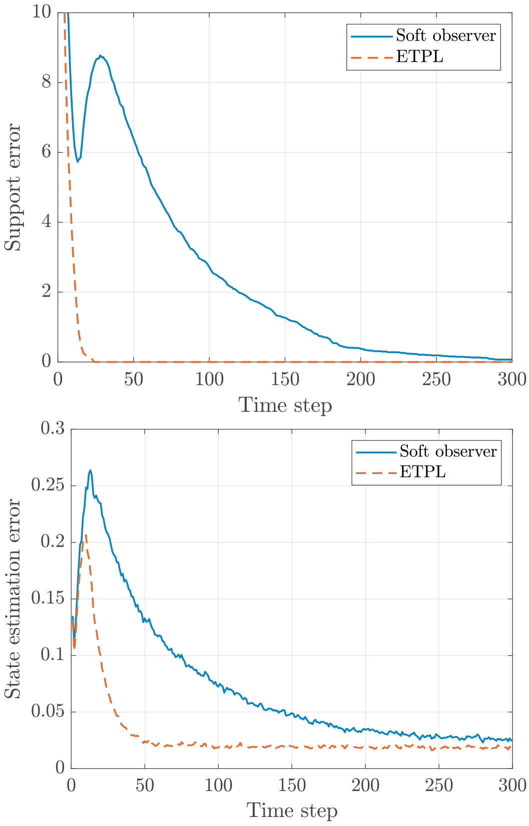

Figure 3. Sparse soft observer vs ETPL; $n = 10$, $p = 15$, $s = 3$, $\tau = n$

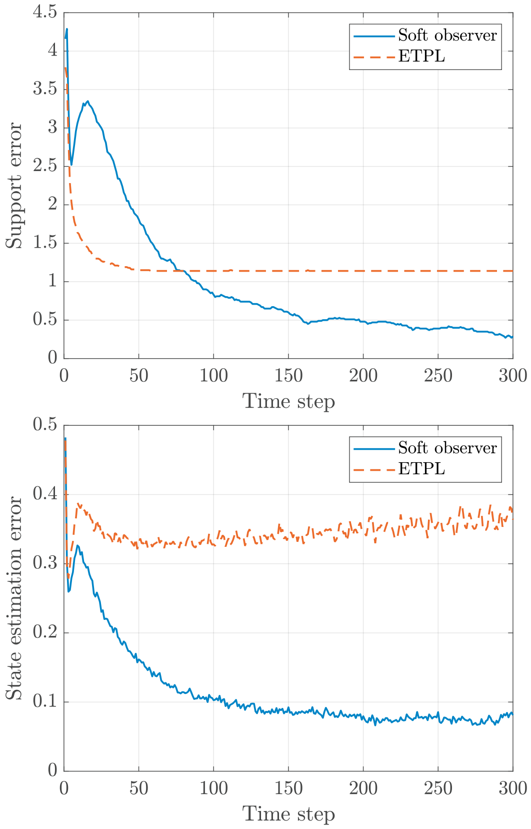

Figure 4. Sparse soft observer vs ETPL; $n = 10$, $p = 15$, $s = 3$, $\tau = 1$

fast time scales. On the other hand, for $\tau = 1$, the sparse observer is more accurate and ETPL does not always converge to the right support. In this case, the execution times are $7 \cdot 10^{-5}$ seconds for ETPL and $4 \cdot 10^{-6}$ seconds for the sparse soft observer.

6. CONCLUSIONS

We propose a Lasso approach for secure state estimation in cyber-physical systems under sparse sensor attacks. We analyse the properties of Lasso to identify the attack and, as a consequence, to recover the state. Furthermore, by starting from the iterative soft thresholding algorithm for Lasso, we develop a sparse soft observer to perform online estimation. Through numerical results, we show that the proposed Lasso approach is valuable with respect to state-of-the-art methods, although it exploits less information, e.g., on the sparsity pattern. Moreover, in our experiments, the sparse soft observer converges to sufficiently accurate solutions, with a reduced execution time. Future work includes the extension of the analysis to noisy models and the study of the convergence of the sparse soft observer.

REFERENCES

Beck, A. and Teboulle, M. (2009). A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.*, 2(1), 183–202.

Cerone, V., Fosson, S.M., and Regruto, D. (2023). Fast sparse optimization via adaptive thresholding. *IFAC-PapersOnLine*, 56(2), 10390–10395.

Cerone, V., Fosson, S.M., Regruto, D., and Salam, A. (2020). Sparse learning with concave regularization: relaxation of the irrepresentable condition. In *IEEE Conf. Decis. Control (CDC)*, 396–401.

Daubechies, I., Defrise, M., and De Mol, C. (2004). An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Comm. Pure Appl. Math.*, 57(11), 1413–1457.

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

Foucart, S. and Rauhut, H. (2013). *A Mathematical Introduction to Compressive Sensing*. Springer, New York.

Fuchs, J.J. (2004). On sparse representations in arbitrary redundant bases. *IEEE Trans. Inf. Theory*, 50(6), 1341–1344.

Fuchs, J.J. (2005). Recovery of exact sparse representations in the presence of bounded noise. *IEEE Trans. Inf. Theory*, 51(10), 3601–3608.

Hastie, T., Tibshirani, R., and Wainwright, M. (2015). *Statistical Learning with Sparsity: The Lasso and Generalizations*. CRC press, 2nd edition.

”Imhotep-SMT” (2015). GitHub repository. URL `http://nesl.github.io/Imhotep-smt/`.

Langner, R. (2011). Stuxnet: Dissecting a cyberwarfare weapon. *IEEE Security & Privacy*, 9(3), 49–51.

Lei, X., Wen, G., Zheng, W.X., and Fu, J. (2023). Security strategy against location-varying sparse attack on distributed state monitoring. *IEEE Trans. Autom. Control*, 1–8.

Lu, A.Y. and Yang, G.H. (2023). A polynomial-time algorithm for the secure state estimation problem under sparse sensor attacks via state decomposition technique. *IEEE Trans. Autom. Control*, 68(12), 7451–7465.

Mao, Y., Mitra, A., Sundaram, S., and Tabuada, P. (2022). On the computational complexity of the secure state-reconstruction problem. *Automatica*, 136, 110083.

Mao, Y. and Tabuada, P. (2023). Decentralized secure state-tracking in multiagent systems. *IEEE Trans. Autom. Control*, 68(7), 4053–4064.

Pajic, M., Lee, I., and Pappas, G.J. (2017). Attack-resilient state estimation for noisy dynamical systems. *IEEE Trans. Control Netw. Syst.*, 4(1), 82–92.

Shoukry, Y., Martin, P., Tabuada, P., and Srivastava, M. (2013). Non-invasive spoofing attacks for anti-lock braking systems. In *CHES*, 55–72.

Shoukry, Y., Nuzzo, P., Puggelli, A., Sangiovanni-Vincentelli, A.L., Seshia, S.A., and Tabuada, P. (2017). Secure state estimation for cyber-physical systems under sensor attacks: A satisfiability modulo theory approach. *IEEE Trans. Autom. Control*, 62(10), 4917–4932.

Shoukry, Y. and Tabuada, P. (2016). Event-triggered state observers for sparse sensor noise/attacks. *IEEE Trans. Autom. Control*, 61(8), 2079–2091.

Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *J. Roy. Stat. Soc. Series B*, 58, 267–288.

Zhao, P. and Yu, B. (2006). On model selection consistency of Lasso. *J. Mach. Learn. Res.*, 7, 2541–2563.

Zhao, Z., Xu, Y., Li, Y., Zhen, Z., Yang, Y., and Shi, Y. (2023). Data-driven attack detection and identification for cyber-physical systems under sparse sensor attacks. *IEEE Trans. Autom. Control*, 68(10), 6330–6337.