Robust and Resilient Estimation for Cyber-Physical Systems under Adversarial Attacks

Sze Zheng Yong [a] Ming Qing Foo [a] Emilio Frazzoli [a]

Abstract— **In this paper, we propose a novel state estimation algorithm that is resilient to *sparse* data injection attacks and robust to *additive* and *multiplicative* modeling errors. By leveraging principles of robust optimization, we construct uncertainty sets that lead to tractable optimization solutions. As a corollary, we obtain a novel robust filtering algorithm when there are no attacks, which can be viewed as a "frequentist" robust estimator as no known priors are assumed. We also describe the use of cross-validation to determine the hyperparameters of our estimator. The effectiveness of our estimator is demonstrated in simulations of an IEEE 14-bus electric power system.**

I. Introduction

Cyber-Physical Systems (CPS) are computer-based systems that monitor and control physical processes using embedded sensors, actuators, control processing units and communication devices. They characterize many of the critical infrastructures that sustain our modern society, such as electric power distribution, oil and natural gas distribution, water and waste-water treatment, and transportation systems. The disruption of these control systems can have disastrous consequences on public health and safety and cause significant economic losses. As CPS become increasingly connected to the internet for remote monitoring and control, they become vulnerable to cyber attacks on their communication channels, which may lead to physical consequences in the forms of faults and failures. Recent incidents of attacks on CPS, including the Maroochy water breach and StuxNet computer worm [1]–[3], emphasize the need for estimation and control algorithms that are resilient to attacks.

Literature Review. Early research on the design of resilient systems has focused on the characterization of undetectable attacks and on attack detection and identification techniques. These range from a simple application of data time-stamps [4] to hypothesis testing using residuals [5]–[8]. More recent works have addressed the problem of state estimation despite attacks, but assume known system parameters and the absence of noise signals. For known linear systems under sparse data injection attack, the resilient state estimation problem is mapped onto an ℓ_0 optimization problem in [9]. The estimator is subsequently relaxed using the "ℓ_1/ℓ_r" norm and demonstrated to be effective under the prescribed conditions. A characterization of the maximum number of correctable attacks was also provided.

For linear systems with bounded additive modeling errors, [10] quantifies the worst-case state estimation error bound when the estimate is generated for the case where the additive modeling errors are implicitly benign, i.e., they can cancel out the attack signals. Hence, this "optimistic" estimate unfortunately does not provide the robustification that we seek. On the other hand, our previous works [11], [12] propose a resilient estimator that generates unbiased estimates asymptotically when the system is perturbed by additive noise signals that are zero mean, Gaussian white processes. Both extensions require the solution of a combinatorial problem, which is intractable for large systems. By contrast, H_∞ filtering approaches (e.g., [13]) cannot take into account the sparse nature of the data injection attack.

Even in the absence of attacks, the robust estimation problem with modeling errors is of significant interest and has been primarily considered from the Bayesian perspective, i.e., with the assumption of known priors. The robust Kalman filtering approach in [14] minimizes the worst-case mean squared state estimation error asymptotically using multiple steady-state Riccati equations, whereas the set-valued filtering approach in [15] utilizes semidefinite relaxation for computing minimal size ellipsoids that bound the solution set of a system of uncertain linear equations.

Another set of relevant literature pertains to that of robust optimization, which addresses the problem of optimization under uncertainty, in which the uncertainty model is not stochastic, but rather deterministic and set-based (e.g., [16], [17]). Of particular relevance is the subject of robust regression and specifically of the equivalence of robustification and regularization in linear regression under some assumptions on the uncertainty sets [18], [19]. This equivalence is a key tool that we will make use of in our estimator design.

Contributions. We propose a novel and computationally tractable state estimation algorithm for uncertain linear systems under adversarial attacks that is resilient to *sparse* data injection attacks (i.e., an adversary can arbitrarily corrupt an unknown but fixed subset of actuators and sensors) and robust to *additive* and *multiplicative* modeling errors. Specifically, we leverage principles of robust optimization and construct uncertainty sets that lead to tractable optimization solutions. As a by-product, we also obtain a novel robust filtering algorithm when there are no attacks, which, in contrast to the existing literature [14], [15], uses a frequentist approach with no assumed known priors. Since it is difficult to predict the modeling errors for the purpose of constructing our uncertainty sets, we use a statistical learning procedure known as cross-validation to determine the hyperparameters of our estimator. As in [9], we illustrative the effectiveness of our approach using random systems and the IEEE 14-bus

[a] S.Z Yong, M.Q. Foo and E. Frazzoli are with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA, USA (e-mail: {szyong, mqf20, frazzoli}@mit.edu).

Preprint submitted to 2016 American Control Conference.
Received September 30, 2015.

electric power system [20].

Notation. For any vector $\mathbf{v} \in \mathbb{R}^n$, $\mathbf{v}_{a:b}$, $1 \leq a \leq b \leq n$, denotes the subset of $\mathbf{v}$ comprising the a-th to b-th entries of $\mathbf{v}$, inclusive. $\mathbf{v}^\top$ denotes the transpose of $\mathbf{v}$. For any matrix $M \in \mathbb{R}^{m\times n}$, $M_{(i,\cdot)} \in \mathbb{R}^n$ denotes the i-th row of M, $i \in \{1, \cdots, m\}$, and $M_{(\cdot,j)} \in \mathbb{R}^m$ denotes the j-th column of M, $j \in \{1, \cdots, n\}$. In addition, the following matrix norms will be used:

- ℓ_0 norm: $\|M\|_{\ell_0}$ = number of nonzero rows of M,
- "mixed" ℓ_1/ℓ_r" norm: $\|M\|_{\ell_1/\ell_r} = \sum_{i=1}^m \|M_{(i,\cdot)}\|_{\ell_r}$,
- (ℓ_q, ℓ_r) subordinate norm: $\|M\|_{(\ell_q,\ell_r)} = \max\limits_{\boldsymbol{\beta}\neq 0} \dfrac{\|M\boldsymbol{\beta}\|_{\ell_r}}{\|\boldsymbol{\beta}\|_{\ell_q}}$.

II. Problem Statement

We model an uncertain CPS that is under attack as the following linear, time invariant (LTI) dynamical system:

$$\begin{aligned} x_{k+1} &= \widetilde{A}\ x_k + \widetilde{B}\ (u_k + d_k) + w_k, \\ y_k &= \widetilde{C}\ x_k + \widetilde{D}\ (u_k + d_k) + e_k + v_k, \end{aligned} \tag{1}$$

where $x_k \in \mathbb{R}^n$ is the state vector at time k, $u_k \in \mathbb{R}^m$ is a known input vector and $y_k \in \mathbb{R}^p$ is the measurement vector. The data injection attacks carried out by the adversary affect the system through the attack signals $d_k \in \mathbb{R}^m$ and $e_k \in \mathbb{R}^p$ that are injected into the actuators and sensors, respectively. The system parameters $\widetilde{A} := A + \Delta A$, $\widetilde{B} := B + \Delta B$, $\widetilde{C} := C + \Delta C$ and $\widetilde{D} := D + \Delta D$ each consist of a known nominal part (A, B, C and D) as well as an unknown part (ΔA, ΔB, ΔC and ΔD) that represents *multiplicative* modeling errors. In addition, the system is affected by *additive* process and measurement modeling errors, $w_k \in \mathbb{R}^n$ and $v_k \in \mathbb{R}^p$, respectively (also referred to as noise signals).

We will assume in this paper that the pair $(\widetilde{A}, \widetilde{C})$ is observable and that the known inputs u_k are independent of x_0 (i.e., we consider the closed loop dynamics in which the dependence of u_k on x_0 is already incorporated into the system). In addition, an adversary attacks a fixed but unknown subset of the sensors and actuators. Note that if sensor $i \in \{1, \cdots, p\}$ is not attacked then necessarily $e_k^{(i)} = 0$ for all time steps k; otherwise $e_k^{(i)}$ can take any value, i.e., the attack signals are sparse, arbitrary and unpredictable. The same observation holds for the attacks on actuators d_k.

In this paper, the term "*resilient*" describes a system that can withstand direct adversarial actions, which we will restrict to data injection attacks [5], [6], [8]. On the other hand, a "*robust*" system can withstand disturbances and modeling errors (also referred to as uncertainties).

The objective of this paper is *robust and resilient state estimation*: Given T corrupted measurements $y_0, y_1, \cdots, y_{T-1}$ and known inputs $u_0, u_1, \cdots, u_{T-1}$, we wish to obtain estimates for the states $x_0, \cdots, x_{T-1}$ that are 1) robust to uncertainties, and 2) resilient to data injection attacks.

III. Preliminary Material

A. Known System with Sensor Attacks Only

We begin with the following simplified system:

$$x_{k+1} = Ax_k, \quad y_k = Cx_k + e_k, \tag{2}$$

where the goal of the estimator is to reconstruct the state sequence $x_0, \cdots, x_{T-1}$ from the corrupted measurements $y_0, \cdots, y_{T-1}$. However, since A is known, the remaining states $x_1, \cdots, x_{T-1}$, can be reconstructed from x_0 using (2) and therefore it is sufficient to recover x_0.

The system (2) can be written compactly as

$$\mathrm{Y} = \Phi(x_0) + \mathrm{E},$$

where $\mathrm{Y} := \begin{bmatrix} y_0 & \cdots & y_{T-1} \end{bmatrix} \in \mathbb{R}^{p\times T}$, $\mathrm{E} := \begin{bmatrix} e_0 & \cdots & e_{T-1} \end{bmatrix} \in \mathbb{R}^{p\times T}$ and Φ is a linear map defined by $\Phi : \mathbb{R}^n \to \mathbb{R}^{p\times T}$:

$$\Phi(x) = \begin{bmatrix} Cx & CAx & \cdots & CA^{T-1}x \end{bmatrix}. \tag{3}$$

The optimal estimator of x_0 for (2) is found by [9] to be

$$x_0 = \underset{x_0\in\mathbb{R}^n}{\arg\min} \|\mathrm{E}\|_{\ell_0} = \underset{x_0\in\mathbb{R}^n}{\arg\min} \|\mathrm{Y} - \Phi(x_0)\|_{\ell_0}. \tag{4}$$

In addition, it was shown that, if (A, C) is observable, then the maximum number of attacked sensors (such that x_0 can be reconstructed exactly) is $\lceil \frac{p}{2} - 1 \rceil$. Moreover, the maximum number of correctable errors cannot increase beyond a window size of $T = n$ measurements (a consequence of Cayley-Hamilton theorem). However, since (4) is intractable (NP-hard), a convex relaxation of the optimal estimator using a "mixed" ℓ_1/ℓ_r norm is considered in [9] that is also used in the compressed sensing literature [21], i.e., the relaxed estimator minimizes the ℓ_1/ℓ_r norm of E:

$$\hat{x}_0 = \underset{x_0\in\mathbb{R}^n}{\arg\min} \|\mathrm{E}\|_{\ell_1/\ell_r} = \underset{x_0\in\mathbb{R}^n}{\arg\min} \|\mathrm{Y} - \Phi(x_0)\|_{\ell_1/\ell_r}. \tag{5}$$

The "hat" on $\hat{x}_0$ denotes that the relaxed estimator (5) generates an estimate of x_0, whereas the optimal estimator (4) recovers the exact x_0. The relaxed estimator (5) has been demonstrated to generate estimates that are close to the exact solutions in [9]. The remaining state estimates $\hat{x}_1, \cdots, \hat{x}_{T-1}$, are then obtained by forward propagation from $\hat{x}_0$ using (2).

B. Known System with Actuator and Sensor Attacks

Next, we consider the following system

$$\begin{aligned} x_{k+1} &= Ax_k + B\,(u_k + d_k), \\ y_k &= Cx_k + D\,(u_k + d_k) + e_k, \end{aligned} \tag{6}$$

which can be written compactly as

$$\mathrm{Y} = \Phi(x_0) + \Theta(\mathrm{U}) + \Theta(\mathrm{D}) + \mathrm{E},$$

where $\mathrm{Y} := \begin{bmatrix} y_0 & \cdots & y_{T-1} \end{bmatrix} \in \mathbb{R}^{p\times T}$, and $\mathrm{U} \in \mathbb{R}^{m\times T}$, $\mathrm{D} \in \mathbb{R}^{m\times T}$ and $\mathrm{E} \in \mathbb{R}^{p\times T}$ are defined similarly. $\Phi : \mathbb{R}^n \to \mathbb{R}^{p\times T}$ is as defined in (3) and $\Theta : \mathbb{R}^{m\times T} \to \mathbb{R}^{p\times T}$ is a linear map defined by

$$\begin{aligned} \Theta(\mathrm{U}) &= \begin{bmatrix} Du_0 & CBu_0 + Du_1 & \cdots & \begin{matrix} C\sum_{i=0}^{T-2} A^{T-2-i}Bu_i \\ +Du_{T-1} \end{matrix} \end{bmatrix} \\ \Theta(\mathrm{D}) &= \begin{bmatrix} Dd_0 & CBd_0 + Dd_1 & \cdots & \begin{matrix} C\sum_{i=0}^{T-2} A^{T-2-i}Bd_i \\ +Dd_{T-1} \end{matrix} \end{bmatrix}. \end{aligned} \tag{7}$$

For (6), the optimal estimator was shown by [9] to be

$$\begin{aligned} (x_0, \mathrm{D}) = \underset{\substack{x_0\in\mathbb{R}^n \\ \mathrm{D}\in\mathbb{R}^{m\times T}}}{\arg\min}\ & \|\mathrm{Y} - \Phi(x_0) - \Theta(\mathrm{U}) - \Theta(\mathrm{D})\|_{\ell_0} \\ & + \|\mathrm{D}\|_{\ell_0}. \end{aligned} \tag{8}$$

In contrast to (4), the optimal estimator in (8) has to generate the initial state x_0 as well as the actuator attacks D so that the remaining states $x_1, \cdots, x_{T-1}$ can be recovered using (6). Similar to (5), the following convex relaxation of (8) is considered:

$$(\hat{x}_0, \hat{\mathrm{D}}) = \underset{\substack{x_0 \in \mathbb{R}^n \\ \mathrm{D} \in \mathbb{R}^{m \times T}}}{\arg\min} \|\mathrm{Y} - \Phi(x_0) - \Theta(\mathrm{U}) - \Theta(\mathrm{D})\|_{\ell_1/\ell_r} + \lambda \|\mathrm{D}\|_{\ell_1/\ell_r} \quad (9)$$

where λ is a tuning parameter. Since the system parameters are known, the estimates of the remaining states $\hat{x}_1, \cdots, \hat{x}_{T-1}$ can be obtained using $\hat{x}_0$, estimates of the actuator attacks $\hat{\mathrm{D}} := \begin{bmatrix} \hat{d}_0 & \cdots & \hat{d}_{T-1} \end{bmatrix}$ and (6).

C. Equivalence of Robust Regression and Regularization

A useful theorem that we shall make use of in our design of a robust estimator is the equivalence of robust regression and regularization under subordinate norm uncertainty sets.

Theorem 1 (Equivalence of Robust Regression and Regularization [19, Corollary 1])**.** *Let $\Delta\Psi$ be an uncertain matrix belonging to the uncertainty set $\mathcal{U}_{(\ell_q,\ell_r)} = \{\Delta\Psi : \|\Delta\Psi\|_{(\ell_q,\ell_r)} \leq \rho\}$. If $q, r \in [1, \infty]$ then for some matrix Ψ and vectors $\mathbf{y}$, $\boldsymbol{\beta}$, we have*

$$\max_{\Delta\Psi \in \mathcal{U}_{(\ell_q,\ell_r)}} \|\mathbf{y} - (\Psi + \Delta\Psi)\boldsymbol{\beta}\|_{\ell_r} = \|\mathbf{y} - \Psi\boldsymbol{\beta}\|_{\ell_r} + \rho\|\boldsymbol{\beta}\|_{\ell_q}.$$

Proof. See [19, Corollary 1]. ■

It is worth noting that there are also similar theorems for the Schatten and Frobenius norms [16], [19], [22].

IV. Robust and Resilient Estimation

Now we are ready to consider the uncertain system under attack in (1), which can be compactly written as

$$\mathrm{Y} = \widetilde{\Phi}(x_0) + \widetilde{\Theta}(\mathrm{U}) + \widetilde{\Theta}(\mathrm{D}) + \widetilde{\Upsilon}(\mathrm{W}, \mathrm{V}) + \mathrm{E} \quad (10)$$

where $\mathrm{Y} := \begin{bmatrix} y_0 & \cdots & y_{T-1} \end{bmatrix} \in \mathbb{R}^{p \times T}$, and $\mathrm{D} \in \mathbb{R}^{m \times T}$, $\mathrm{U} \in \mathbb{R}^{p \times T}$, $\mathrm{W} \in \mathbb{R}^{n \times T}$, $\mathrm{V} \in \mathbb{R}^{p \times T}$ and $\mathrm{E} \in \mathbb{R}^{p \times n}$ are defined similarly. $\widetilde{\Phi}$, $\widetilde{\Theta}$ and $\widetilde{\Upsilon}$ are linear maps, with $\widetilde{\Phi} : \mathbb{R}^n \to \mathbb{R}^{p \times T}$ and $\widetilde{\Theta} : \mathbb{R}^{m \times T} \to \mathbb{R}^{p \times T}$ that are similarly defined as in (3) and (7), except with the true matrices, $\widetilde{A}$, $\widetilde{B}$, $\widetilde{C}$ and $\widetilde{D}$, while $\widetilde{\Upsilon} : \mathbb{R}^{n \times T} \times \mathbb{R}^{p \times T} \to \mathbb{R}^{p \times T}$ is:

$$\widetilde{\Upsilon}(\mathrm{W}, \mathrm{V}) = \begin{bmatrix} v_0 & \widetilde{C} w_0 + v_1 & \cdots & \begin{matrix} \widetilde{C} \sum_{i=0}^{T-2} \widetilde{A}^{T-2-i} w_i \\ + v_{T-1} \end{matrix} \end{bmatrix}. \quad (11)$$

In light of the uncertain parameters in (1), we consider the robustification of the estimator in (9) by using the compact representation in (10), i.e.,

$$\begin{aligned} (\hat{x}_0, \hat{\mathrm{D}}) &= \underset{\substack{x_0 \in \mathbb{R}^n \\ \mathrm{D} \in \mathbb{R}^{m \times T}}}{\arg\min} \max_{\Delta\Psi \in \mathcal{U}_{(\ell_q,\ell_r)}} \|\mathrm{E}\|_{\ell_1/\ell_r} + \lambda \|\mathrm{D}\|_{\ell_1/\ell_r} \\ &= \underset{\substack{x_0 \in \mathbb{R}^n \\ \mathrm{D} \in \mathbb{R}^{m \times T}}}{\arg\min} \max_{\Delta\Psi \in \mathcal{U}_{(\ell_q,\ell_r)}} \left\| \begin{matrix} \mathrm{Y} - \widetilde{\Phi}(x_0) - \widetilde{\Theta}(\mathrm{U}) \\ - \widetilde{\Theta}(\mathrm{D}) - \widetilde{\Upsilon}(\mathrm{W}, \mathrm{V}) \end{matrix} \right\|_{\ell_1/\ell_r} \\ &\qquad + \lambda \|\mathrm{D}\|_{\ell_1/\ell_r} \end{aligned} \quad (12)$$

for some tuning parameter λ and some uncertain $\Delta\Psi$ belonging to the uncertainty set $\mathcal{U}_{(\ell_q,\ell_r)}$, which will be described in Definition 1. In Section IV-A, we will provide $\Delta\Psi$ and $\mathcal{U}_{(\ell_q,\ell_r)}$ that will lead to tractable formulations of (12).

A. Robust and Resilient Estimation of Initial State x_0 and Actuator Attack Signals D

Similar to the approach in [9], we first find estimates of x_0 and the actuator attacks D. Notice that it is helpful to consider another compact representation of (1):

$$\mathbf{y} = \widetilde{\mathcal{O}} x_0 + \widetilde{\mathcal{J}}_u(\mathbf{u} + \mathbf{d}) + \widetilde{\mathcal{J}}_w \mathbf{w} + \mathbf{e} + \mathbf{v} \quad (13)$$

with $\mathbf{y} := \mathrm{vec}(\mathrm{Y})$, $\mathbf{u} := \mathrm{vec}(\mathrm{U})$, $\mathbf{d} := \mathrm{vec}(\mathrm{D})$, $\mathbf{e} := \mathrm{vec}(\mathrm{E})$, $\mathbf{w} := \mathrm{vec}(\mathrm{W})$ and $\mathbf{v} := \mathrm{vec}(\mathrm{V})$, where $\mathrm{vec}(\cdot)$ is the vectorization operator, and we also define the following observability and invertibility matrices

$$\widetilde{\mathcal{O}} = \begin{bmatrix} \widetilde{C}^\top & (\widetilde{C}\widetilde{A})^\top & (\widetilde{C}\widetilde{A}^2)^\top & \ldots & (\widetilde{C}\widetilde{A}^{T-1})^\top \end{bmatrix}^\top,$$

$$\widetilde{\mathcal{J}}_u = \begin{bmatrix} \widetilde{D} & 0 & 0 & \ldots & 0 \\ \widetilde{C}\widetilde{B} & \widetilde{D} & 0 & \ldots & 0 \\ \widetilde{C}\widetilde{A}\widetilde{B} & \widetilde{C}\widetilde{B} & \widetilde{D} & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \widetilde{C}\widetilde{A}^{T-2}\widetilde{B} & \widetilde{C}\widetilde{A}^{T-3}\widetilde{B} & \widetilde{C}\widetilde{A}^{T-4}\widetilde{B} & \cdots & \widetilde{D} \end{bmatrix},$$

$$\widetilde{\mathcal{J}}_w = \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 \\ \widetilde{C} & 0 & 0 & \ldots & 0 \\ \widetilde{C}\widetilde{A} & \widetilde{C} & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \widetilde{C}\widetilde{A}^{T-2} & \widetilde{C}\widetilde{A}^{T-3} & \widetilde{C}\widetilde{A}^{T-4} & \cdots & 0 \end{bmatrix}.$$

The matrices $\mathcal{O}$ and $\mathcal{J}_u$ are defined in a similar fashion with the nominal system matrices A, B, C and D. We also define $\Delta\mathcal{O} := \widetilde{\mathcal{O}} - \mathcal{O}$ and $\Delta\mathcal{J}_u := \widetilde{\mathcal{J}}_u - \mathcal{J}_u$.

1) Row-wise Uncertainty Sets: We now define an uncertainty set $\mathcal{U}_{(\ell_q,\ell_r)}$ that captures the notion of boundedness of both the additive and multiplicative modeling errors. Specifically, we let $\Delta\Psi \in \mathcal{U}_{(\ell_q,\ell_r)}$ represent all row-wise uncertainty sets $\Delta\Psi_i \in \mathcal{U}_{i,(\ell_q,\ell_r)}$ for $i = 1, \cdots, p$, which we assume are uncoupled from each other and are defined as follows:

Definition 1 (Row-wise uncertainty sets)**.** *For each $i = 1, \cdots, p$, we define the row-wise uncertainty matrix as*

$$\Delta\Psi_i := \begin{bmatrix} (\Delta\mathcal{O})_i & (\Delta\mathcal{J}_u)_i & (\Delta\mathcal{J}_u)_i & (\widetilde{\mathcal{J}}_w)_i \mathbf{w} + (\mathbf{v})_i \end{bmatrix},$$

which we assume belong to the uncertainty set $\mathcal{U}_{i,(\ell_q,\ell_r)} = \{\Delta\Psi_i : \|\Delta\Psi_i\|_{(\ell_q,\ell_r)} \leq \rho_i\}$, with $(M)_i$ denoting the submatrix of M consisting of only the $(i + jp)$-th rows of M for $j = 0, \cdots, T-1$. (e.g., $(\widetilde{\mathcal{O}})_i := \begin{bmatrix} \widetilde{C}_{(i,\cdot)} & (\widetilde{C}\widetilde{A})_{(i,\cdot)} & (\widetilde{C}\widetilde{A}^2)_{(i,\cdot)} & \ldots & (\widetilde{C}\widetilde{A}^{T-1})_{(i,\cdot)} \end{bmatrix}^\top$).

2) Estimator Design: Notice that with the definition of the ℓ_1/ℓ_r norm for E, i.e., $\|\mathrm{E}\|_{\ell_1/\ell_r} = \sum_{i=1}^p \|\mathrm{E}_{(i,\cdot)}\|_{\ell_r}$ and the assumption of uncoupled row-wise uncertainty sets as defined in Definition 1, we can consider the problem in (12) by first considering a collection of subproblems for each of the i-th row of E, i.e.,

$$\begin{aligned} &\max_{\Delta\Psi_i \in \mathcal{U}_{i,(\ell_q,\ell_r)}} \|\mathrm{E}_{(i,\cdot)}\|_{\ell_r} \\ &= \max_{\Delta\Psi_i \in \mathcal{U}_{i,(\ell_q,\ell_r)}} \left\| \begin{matrix} (\mathrm{Y} - \widetilde{\Phi}(x_0) - \widetilde{\Theta}(\mathrm{U}) \\ - \widetilde{\Theta}(\mathrm{D}) - \widetilde{\Upsilon}(\mathrm{W}, \mathrm{V}))_{(i,\cdot)} \end{matrix} \right\|_{\ell_r} \end{aligned} \quad (14)$$

for some uncertain matrix $\Delta\Psi_i$ belonging to the row-wise uncertainty set $\mathcal{U}_{i,(\ell_q,\ell_r)}$ as in Definition 1. The subproblems can be simplified as is shown in the following lemma.

Lemma 1. *Let $\Delta\Psi_i$ and $\mathcal{U}_{i,(\ell_q,\ell_r)}$ be defined according to Definition 1. In addition, we define*

$$\begin{aligned}\Psi_i &:= \begin{bmatrix}(\mathcal{O})_i & (\mathcal{J}_u)_i & (\mathcal{J}_u)_i & \mathbf{0}_{T\times 1}\end{bmatrix},\\ \boldsymbol{\beta} &:= \begin{bmatrix}x_0^\top & \mathbf{u}^\top & \mathbf{d}^\top & 1\end{bmatrix}^\top.\end{aligned}$$

Then, for any $q, r \in [1,\infty]$, (14) is equivalent to

$$\max_{\Delta\Psi_i \in \mathcal{U}_{i,(\ell_q,\ell_r)}} \left\|\mathrm{E}_{(i,\cdot)}\right\|_{\ell_r} = \|\mathrm{Y}_{(i,\cdot)} - \Psi_i\boldsymbol{\beta}\|_{\ell_r} + \rho_i\|\boldsymbol{\beta}\|_{\ell_q}.$$

Proof. We can rewrite $\mathrm{E}_{(i,\cdot)}$ as follows

$$\begin{aligned}\mathrm{E}_{(i,\cdot)} &= \left(\mathrm{Y} - \widetilde{\Phi}(x_0) - \widetilde{\Theta}(\mathrm{U}) - \widetilde{\Theta}(\mathrm{D}) - \widetilde{\Upsilon}(\mathrm{W},\mathrm{V})\right)_{(i,\cdot)}\\ &= \mathrm{Y}_{(i,\cdot)} - (\Psi_i + \Delta\Psi_i)\boldsymbol{\beta}.\end{aligned}$$

The result follows by applying Theorem 1 to (14). ∎

Now we are ready to develop a robust estimator for (12), followed by a corollary for the special case of $\ell_q = \ell_1$ [1].

Proposition 1 (Robust Estimation of x_0 with ℓ_1/ℓ_r relaxation and ℓ_q-regularization)**.** *Let $\Delta\Psi \in \mathcal{U}_{(\ell_q,\ell_r)}$ represent uncoupled $\Delta\Psi_i \in \mathcal{U}_{i,(\ell_q,\ell_r)}$ for all $i = 1,\cdots,p$. Then, for any $q, r \in [1,\infty)$, the robust estimator is equivalent to the following constrained optimization problem*

$$\begin{aligned}(\hat{x}_0, \hat{\mathrm{D}}) \;=\; \mathop{\arg\min}_{\substack{x_0\in\mathbb{R}^n\\ \mathrm{D}\in\mathbb{R}^{m\times T}\\ \boldsymbol{\beta}\in\mathbb{R}^{n+2mT+1}}} \;& \|\mathrm{Y} - \Phi(x_0) - \Theta(\mathrm{U}) - \Theta(\mathrm{D})\|_{\ell_1/\ell_r}\\ &+\lambda\|\mathrm{D}\|_{\ell_1/\ell_r} + \rho\|\boldsymbol{\beta}\|_{\ell_q}\\ \textit{s.t.} \quad & \boldsymbol{\beta}_{1:n} = x_0\\ & \boldsymbol{\beta}_{n+1:n+mT} = \mathrm{vec}(\mathrm{U})\\ & \boldsymbol{\beta}_{n+mT+1:n+2mT} = \mathrm{vec}(\mathrm{D})\\ & \boldsymbol{\beta}_{n+2mT+1} = 1\end{aligned}$$

with $\rho = \sum_{i=1}^{p}\rho_i$ (cf. Definition 1). $\lambda > 0$ is a tuning parameter that controls the relative weight between penalties on errors corresponding to attacks on sensors and actuators.

Proof. This proposition follows the repeated application of Lemma 1 and by noticing that $Y_{(i,\cdot)} - \Psi_i\boldsymbol{\beta} = (\mathrm{Y} - \Phi(x_0) - \Theta\,(\mathrm{U}) - \Theta(\mathrm{D}))_{(i,\cdot)}$. ∎

Corollary 1 (Robust Estimation of x_0 with ℓ_1/ℓ_r relaxation and ℓ_1-regularization)**.** *Let $\Delta\Psi \in \mathcal{U}_{(\ell_1,\ell_r)}$ represent $\Delta\Psi_i \in \mathcal{U}_{i,(\ell_1,\ell_r)}$ for all $i = 1,\cdots,p$. Then, for any $r \in [1,\infty]$,*

$$\begin{aligned}(\hat{x}_0, \hat{\mathrm{D}}) = \mathop{\arg\min}_{\substack{x_0\in\mathbb{R}^n\\ \mathrm{D}\in\mathbb{R}^{m\times T}}} \quad & \|\mathrm{Y} - \Phi(x_0) - \Theta(\mathrm{U}) - \Theta(\mathrm{D})\|_{\ell_1/\ell_r}\\ & +\rho\|x_0\|_{\ell_1} + \rho\|\mathrm{D}\|_{\ell_1/\ell_1} + \lambda\|\mathrm{D}\|_{\ell_1/\ell_r}.\end{aligned}$$

with $\rho = \sum_{i=1}^{p}\rho_i$ (cf. Definition 1). $\lambda > 0$ is a tuning parameter that controls the relative weight between penalties on errors corresponding to attacks on sensors and actuators.

Proof. Noting that $\|\boldsymbol{\beta}\|_{\ell_1} = \|x_0\|_{\ell_1} + \|\mathbf{u}\|_{\ell_1} + \|\mathbf{d}\|_{\ell_1} + 1$, $\|\mathbf{d}\|_{\ell_1} = \|\mathrm{D}\|_{\ell_1/\ell_1}$ and $Y_{(i,\cdot)} - \Psi_i\boldsymbol{\beta} = (\mathrm{Y} - \Phi(x_0) - \Theta\,(\mathrm{U}) - \Theta(\mathrm{D}))_{(i,\cdot)}$, the application of Lemma 1 on (14) gives

$$\begin{aligned}\max_{\Delta\Psi_i \in \mathcal{U}_{i,(\ell_q,\ell_r)}} \left\|\mathrm{E}_{(i,\cdot)}\right\|_{\ell_r} = & \left\|(\mathrm{Y} - \Phi(x_0) - \Theta(\mathrm{U}) - \Theta(\mathrm{D}))_{(i,\cdot)}\right\|_{\ell_r}\\ & + \rho_i(\|x_0\|_{\ell_1} + \|\mathbf{u}\|_{\ell_1} + \|\mathbf{d}\|_{\ell_1} + 1).\end{aligned}$$

Since we have assumed that $\mathbf{u}$ is independent of x_0, we can use the above for each i to obtain the equivalence of (12) to the expression in the corollary statement. ∎

3) Summary: A key insight we gained is that, with an appropriate choice of an uncertainty set from Definition 1, a robustification of (9) is equivalent to a regularization procedure. Note that we denote the resulting robust and resilient estimates of Proposition 1 as $(\hat{x}_{0,\mathrm{rob}}^{\ell_1/\ell_r}, \hat{\mathrm{D}}_{\mathrm{rob}})$, and the estimates of the nominal estimator in (9) as $(\hat{x}_{0,\mathrm{nom}}^{\ell_1/\ell_r}, \hat{\mathrm{D}}_{\mathrm{nom}})$. In contrast to the nominal estimator, the robust and resilient estimator has an additional parameter ρ that controls the amount of robustification (a greater ρ indicates a more conservative estimator). Moreover, the nominal estimator is equivalent to the robust version with $\rho = 0$.

Remark 1. *In practice, it is difficult to construct ρ because the modeling errors cannot be accurately predicted. In addition, there is no clear strategy for selecting the ideal values for λ, ℓ_q and ℓ_r. Thus, it is natural to use a statistical approach such as cross-validation with data sets to obtain these hyperparameters, which will be discussed in detail in Section V-B.1.*

B. Robust Estimation of the Remaining States

Even with the estimates $(\hat{x}_0, \hat{\mathrm{D}})$, we cannot obtain $\hat{x}_1,\cdots,\hat{x}_{T-1}$ using (1) because the system parameters $\widetilde{A}, \widetilde{B}$ and the additive error w_k are unknown. Therefore, we now develop a robust estimator for the states $x_1,\cdots,x_{T-1}$ using $(\hat{x}_0, \hat{\mathrm{D}})$. This second problem can be formulated as: Given $(\hat{x}_0, \hat{\mathrm{D}})$ and known inputs $\mathbf{u}$, we wish to obtain estimates of the remaining states $X := \begin{bmatrix}x_1^\top & x_2^\top & \cdots & x_{T-1}^\top\end{bmatrix}^\top$ that are robust to modeling errors ΔA, ΔB and w_k.

First, note that state equations in (1) can be written as

$$X = \widetilde{\mathcal{P}}x_0 + \widetilde{\mathcal{K}}_u(\mathbf{u} + \mathbf{d}) + \widetilde{\mathcal{K}}_w\mathbf{w}, \tag{15}$$

where the state transition and input matrices are given by

$$\begin{aligned}\widetilde{\mathcal{P}} &= \begin{bmatrix}(\widetilde{A})^\top & (\widetilde{A}^2)^\top & \ldots & (\widetilde{A}^{T-1})^\top\end{bmatrix}^\top,\\ \widetilde{\mathcal{K}}_u &= \begin{bmatrix}\widetilde{B} & 0 & 0 & \ldots & 0 & 0\\ \widetilde{A}\widetilde{B} & \widetilde{B} & 0 & \ldots & 0 & 0\\ \widetilde{A}^2\widetilde{B} & \widetilde{A}\widetilde{B} & \widetilde{B} & \ldots & 0 & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ \widetilde{A}^{T-2}\widetilde{B} & \widetilde{A}^{T-3}\widetilde{B} & \widetilde{A}^{T-4}\widetilde{B} & \cdots & \widetilde{B} & 0\end{bmatrix},\\ \widetilde{\mathcal{K}}_w &= \begin{bmatrix}I & 0 & 0 & \ldots & 0 & 0\\ \widetilde{A} & I & 0 & \ldots & 0 & 0\\ \widetilde{A}^2 & \widetilde{A} & I & \ldots & 0 & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ \widetilde{A}^{T-2} & \widetilde{A}^{T-3} & \widetilde{A}^{T-4} & \cdots & I & 0\end{bmatrix}.\end{aligned} \tag{16}$$

The matrices $\mathcal{P}$, $\mathcal{K}_u$ and $\mathcal{K}_w$ are similarly defined with A and B instead of $\widetilde{A}$ and $\widetilde{B}$. In addition, we define $\Delta\mathcal{P} :=$

[1] Besides simplifying the robust estimator, this choice is also observed to be justified in simulations in Section V-B.5.

$\widetilde{\mathcal{P}}-\mathcal{P}$ and $\Delta\mathcal{K}_u := \widetilde{\mathcal{K}}_u-\mathcal{K}_u$. In the following, we show that the robust estimates can be essentially obtained by forward propagation using the nominal state dynamics, i.e., $\hat{x}_{k+1} = A\hat{x}_k + B(u_k+\hat{d}_k)$ with given $(\hat{x}_0, \hat{\mathrm{D}})$ and $\hat{\mathbf{d}} := \mathrm{vec}(\hat{\mathrm{D}})$.

Definition 2. *Let* $\Delta\Omega := \begin{bmatrix}\Delta\mathcal{P} & \Delta\mathcal{K}_u & \widetilde{\mathcal{K}}_w\mathbf{w}\end{bmatrix}$ *be an uncertain matrix belonging to the uncertainty set* $\mathcal{U}_{(\ell_q,\ell_r)} = \{\Delta\Omega : \|\Delta\Omega\|_{(\ell_q,\ell_r)} \le \tilde{\rho}\}$.

Proposition 2 (Robust Estimation of State Sequence)**.** *Let* $\Delta\Omega$ *and* $\mathcal{U}_{(\ell_q,\ell_r)}$ *be defined according to Definition 2. Then, given* $x_0 = \hat{x}_0$, $\mathbf{d} = \hat{\mathbf{d}} = \mathrm{vec}\Big(\hat{\mathrm{D}}\Big)$ *and* $\mathbf{u}$, *the robust estimate of* X *for any* $q, r \in [1, \infty]$ *is given by*

$$\begin{aligned}\hat{X} &= \underset{X\in\mathbf{R}^{n(T-1)}}{\arg\min}\ \max_{\Delta\Omega\in\mathcal{U}_{(\ell_q,\ell_r)}} \|X-\widetilde{\mathcal{P}}\hat{x}_0-\widetilde{\mathcal{K}}_u(\mathbf{u}+\hat{\mathbf{d}})-\widetilde{\mathcal{K}}_w\mathbf{w}\|_{\ell_r}\\ &= \mathcal{P}\hat{x}_0+\mathcal{K}_u(\mathbf{u}+\hat{\mathbf{d}}).\end{aligned}$$

Proof. From (15) and the definitions in (16),

$$X-\widetilde{\mathcal{P}}\hat{x}_0-\widetilde{\mathcal{K}}_u(\mathbf{u}-\hat{\mathbf{d}})-\widetilde{\mathcal{K}}_w\mathbf{w} = X-(\Omega+\Delta\Omega)\boldsymbol{\gamma},$$

where $\Omega := \begin{bmatrix}\mathcal{P} & \mathcal{K}_u & 0\end{bmatrix}$, $\Delta\Omega := \begin{bmatrix}\Delta\mathcal{P} & \widetilde{\mathcal{K}}_u & \widetilde{\mathcal{K}}_w\mathbf{w}\end{bmatrix}$ and $\boldsymbol{\gamma} = \begin{bmatrix}\hat{x}_0^\top & \mathbf{u}^\top+\hat{\mathbf{d}}^\top & 1\end{bmatrix}^\top$. Then, by Theorem 1, we have

$$\begin{aligned}\hat{X} &= \underset{X\in\mathbf{R}^{n(T-1)}}{\arg\min}\ \max_{\Delta\Omega\in\mathcal{U}_{(\ell_q,\ell_r)}} \|X-(\Omega+\Delta\Omega)\boldsymbol{\gamma}\|_{\ell_r}\\ &= \underset{X\in\mathbf{R}^{n(T-1)}}{\arg\min}\ \|X-\Omega\boldsymbol{\gamma}\|_{\ell_r}+\tilde{\rho}\|\boldsymbol{\gamma}\|_{\ell_q}\\ &= \underset{X\in\mathbf{R}^{n(T-1)}}{\arg\min}\ \|X-\mathcal{P}\hat{x}_0-\mathcal{K}_u(\mathbf{u}+\hat{\mathbf{d}})\|_{\ell_r}+\tilde{\rho}\|\boldsymbol{\gamma}\|_{\ell_q}\\ &= \mathcal{P}\hat{x}_0+\mathcal{K}_u(\mathbf{u}+\hat{\mathbf{d}})\end{aligned}$$

since $\boldsymbol{\gamma}$ is known and $\|z\|_{\ell_r} = 0$ if and only if $z = 0$. ■

C. Robust Filtering as a Corollary

In the absence of attacks (cf. Scenario (ii) in Section V-B.3), the robust and resilient estimator developed in the previous sections provides an alternative approach to robust filtering for the following system

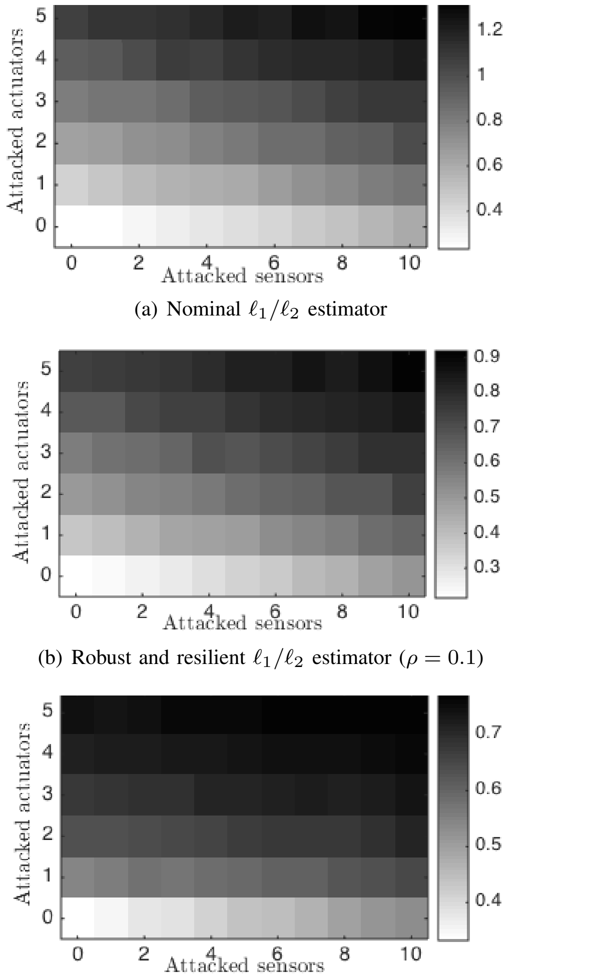
$$\begin{aligned}x_{k+1} &= \widetilde{A}\ x_k+\widetilde{B}\ u_k+w_k\\ y_k &= \widetilde{C}\ x_k+\widetilde{D}\ u_k+v_k.\end{aligned}$$

The goal of our robust filter is: Given T noisy measurements $y_1, \cdots, y_{T-1}$ and known inputs $u_1, \cdots, u_{T-1}$, we wish to estimate the states $x_0, \cdots, x_{T-1}$. Due to space limitations, details of its derivation will be provided in a later publication. However, it is not far-fetched to see that the robust and resilient estimator we developed in Sections IV-A and IV-B is also applicable as a robust filter even in the absence of attacks.

V. Numerical Simulations

A. Random systems

We first compare performances of the ℓ_1/ℓ_2 nominal resilient estimator in (9) and the ℓ_1/ℓ_2 robust and resilient estimator in Section IV in predicting the randomly chosen initial state x_0 (from a Gaussian distribution with variance 1) on random systems described by (1) of size $n = 20$ states, $m = 5$ actuators and $p = 20$ sensors. The matrices $\{B, C, D\}$, $\{A, \Delta B, \Delta C, \Delta D\}$ and ΔA have i.i.d. Gaussian entries with variances 1, 0.01 and 10^{-4}, respectively. For different numbers of attacked actuators and sensors, we tested the estimators on 500 random systems with a window size of $T = 30$, random initial conditions and randomly chosen sets of attacked actuators and sensors. The resulting normalized estimation error for the initial state x_0, averaged over the 500 instances are shown in Figure 1. The results indicate that the robust and resilient estimators ($\rho = \{0.1, 1\}$ and $\ell_q = \ell_1$) consistently perform better than the nominal estimator. On the other hand, a higher ρ value decreases the range of normalized errors and the largest mean value, but with a slightly higher smallest mean value (compare the ranges of the color bars in Figures 1(b) and 1(c)).



(a) Nominal ℓ_1/ℓ_2 estimator

(b) Robust and resilient ℓ_1/ℓ_2 estimator ($\rho = 0.1$)

(c) Robust and resilient ℓ_1/ℓ_∞ estimator ($\rho = 1$)

Fig. 1. Mean of normalized errors of the nominal and robust ℓ_1/ℓ_2 estimators (with $\lambda = 0.2$, $\ell_q = \ell_1$) averaged over 500 random systems for varying numbers of attacked actuators and sensors. A darker shade (note the different color bar scalings) indicates a higher mean of normalized errors.

B. Electric Power Network

We now demonstrate the effectiveness of our robust and resilient estimator in Proposition 1 over the nominal estimator (9) using an IEEE 14-bus system [20], [23] that is subject to data injection attacks. The system, depicted in Figure 2, consists of 5 synchronous generators and 14 buses. It is

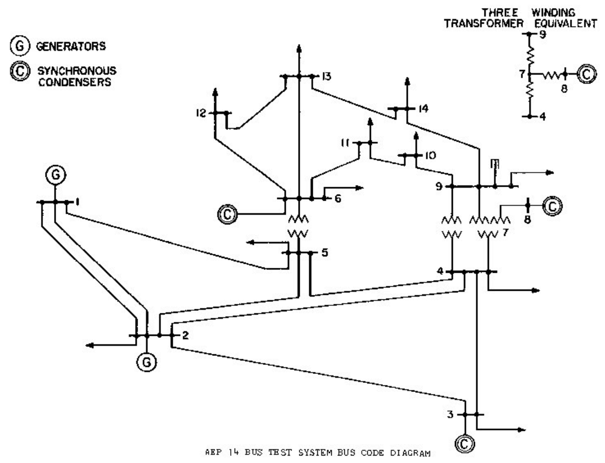

G GENERATORS
C SYNCHRONOUS CONDENSERS
THREE WINDING TRANSFORMER EQUIVALENT
1
2
3
4
5
6
7
8
9
10
11
12
13
14
AEP 14 BUS TEST SYSTEM BUS CODE DIAGRAM

Fig. 2. IEEE 14-bus electric power system [20]

represented by $n = 10$ states comprising the rotor angles and frequencies of each generator. The dynamics of the system can be represented by an uncertain LTI model (see [24] for the derivation of the linearized swing equations), that is discretized with a sampling interval of $dT = 0.05s$ to obtain the model in (1) with $\tilde{D} = 0$. Similar to [23], $p = 35$ sensors are deployed to measure the real power injections at every bus, the real power flows along every branch and the rotor angle of generator 1, with the sensor measuring the rotor angle of generator 1 being protected from attacks.

In the following sections, we will describe the cross validation procedure used to determine the hyperparameters of our robust and resilient estimator, as well as the simulations used to validate the effectiveness of our robust and resilient estimator in Proposition 1 over the nominal estimator in (9). The estimators are implemented in MATLAB and a MATLAB interface to CVX [25], [26] is used to solve the optimization problems. In all our simulations, the initial state $x(0) = x_0$ and modeling errors ($\Delta A, \Delta B, \Delta C$, w_k and v_k) are drawn from i.i.d. Gaussian distributions, while the nominal matrices are as in [23].

1) Cross Validation Procedure for Selection of Hyperparameters: In practice, it is difficult to predict the modeling errors and noise signals for the purpose of constructing our uncertainty sets. Thus, it is natural to use a statistical learning procedure known as *cross-validation* to determine the hyperparameters of our robust and resilient estimator – namely, given some training data, we want to select i) the tuning parameter λ, ii) the robustification level ρ, and iii) the estimation approach among our robust and resilient ℓ_1/ℓ_1, ℓ_1/ℓ_2 and ℓ_1/ℓ_∞ estimators (with $\ell_q = \ell_1$).

To this end, 200 sets of data (given by the tuple $(x_0, y_0, y_1, \ldots, y_{T-1})$ are generated with a window size of $T = 15$) using a nominal system model with modeling errors and attack signals drawn from i.i.d. Gaussian distributions, initial states x_0 drawn from the standard Gaussian distribution, different sets of attacked sensors K of cardinality $q_s = 3$ and different sets of attacked actuators L of cardinality $q_a = 1$. Subsequently, the data is randomly partitioned into three sets: allocate 50% for *training*, 25% for *validiation* and 25% for *testing*. The procedure of cross-validation for both the nominal and robust resilient estimators is conducted in the following phases:

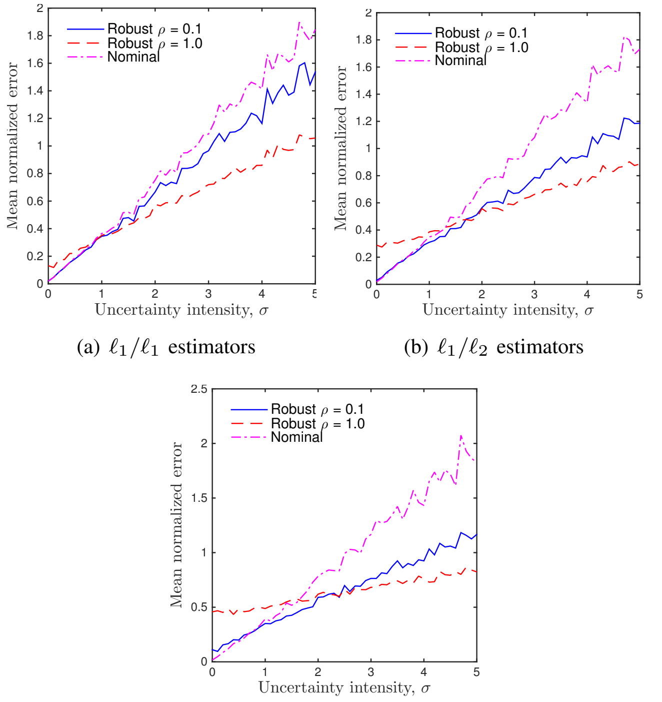
Training: For each approach (ℓ_1/ℓ_1, ℓ_1/ℓ_2 and ℓ_1/ℓ_∞), find the best values of λ and ρ for the *training* set.

Validation: Using the *validation* set, select the best approach among the ℓ_1/ℓ_1, ℓ_1/ℓ_2 and ℓ_1/ℓ_∞ estimators with λ and ρ determined in the training phase.

Testing: Determine how well the resilient estimator (nominal and robust) can predict the values of x_0 in the *testing* set.

When the above process is repeated 20 times, a comparison of initial state estimation errors of the robust and resilient estimator to the estimation errors of the nominal estimator shows average reductions of 16.92% and 11.68% in the mean and standard deviation, respectively. Furthermore, when the intensities of the model errors (i.e., the variance of the i.i.d. Gaussian distribution from which the error samples are drawn) are increased by about 2.5 times, a similar cross-validation study shows decreases of 14.06% and 41.43% in the mean and standard deviation, respectively, in the initial state estimation errors.

2) Varying intensities of modeling errors: To observe the effects of modeling errors on the performances of our estimators, their intensities (i.e., variances of the i.i.d. zero-mean Gaussian perturbations) are varied while the parameters λ and ρ are kept constant. For different intensity levels, the simulations are repeated 100 times with different sets of attacked sensors K of cardinality $q_s = 3$ and different sets of attacked actuators L of cardinality $q_a = 1$.



(a) ℓ_1/ℓ_1 estimators

(b) ℓ_1/ℓ_2 estimators

(c) ℓ_1/ℓ_∞ estimators

Fig. 3. Mean normalized errors of the nominal estimator and the robust and resilient ℓ_1/ℓ_1, ℓ_1/ℓ_2 and ℓ_1/ℓ_∞ estimators (with $\ell_q = \ell_1$, $\lambda = 0.2$ and different values for ρ) simulated on the IEEE 14-bus system.

The procedure is repeated for different robust and resilient estimators (ℓ_1/ℓ_1, ℓ_1/ℓ_2 and ℓ_1/ℓ_∞ with $\ell_q = \ell_1$) and compared with the nominal estimators. From Figure 3, it is clear that the nominal estimators perform the best when the uncertainty intensity is small. But, as the uncertainty intensity is increased, the situation is reversed, and a larger ρ (i.e., more conservative robust and resilient estimator) leads to better estimates.

3) Estimation of the Initial State for Different Attack and Uncertainty Scenarios: Next, we compare the performances of the ℓ_1/ℓ_∞ robust and resilient estimator and the ℓ_1/ℓ_∞ nominal estimator (with $\ell_q = \ell_1$) for various scenarios: (i) "*attack only*" (modeling errors are absent), (ii) "*uncertainty only*" (attack signals are absent) and (iii) "*uncertainty and attack*" (modeling errors and attacks are present).

The results of 500 simulations are summarized in Figure 4. In the "*attack only*" scenario, it can be observed that the nominal estimator performs the best, thus confirming our observations of Section V-B.2 and validating the effectiveness of the nominal estimator for known systems in [9]. In the "*uncertainty only*" scenario (i.e., the robust filtering scenario mentioned in Section IV-C), a significant improvement in performance of the robust and resilient estimator over the nominal estimator can be observed. The same observations can be seen in the "*uncertainty and attack*" scenario. Thus, we can conclude that our robust and resilient estimator is more effective than the nominal estimator for estimating the initial state when there are modeling errors.

4) Estimation of the State Trajectory for Different Attack and Uncertainty Scenarios: Now that we have obtained the initial state estimates in the previous section, we proceed to estimate the remaining states of the state trajectory, using the approach we developed in Section IV-B. The results of 500 simulations are summarized in Figure 5. As expected, the nominal estimator performs best in the "*attack only*" scenario (practically zero for all times), but fares worse than the robust and resilient estimators in the other scenarios.

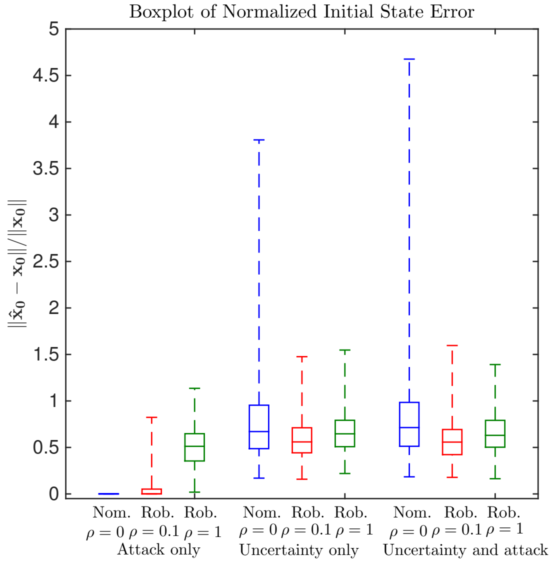


Fig. 4. Normalized errors of the nominal ℓ_1/ℓ_∞ estimator and the robust and resilient ℓ_1/ℓ_∞ estimator (with $\ell_q = \ell_1$, $\lambda = 0.2$ and different values for ρ) under different scenarios simulated on the IEEE 14-bus system. The dashed lines represent the support of the data, while the box represents the median as well as the 25th and 75th percentiles of the normalized errors.

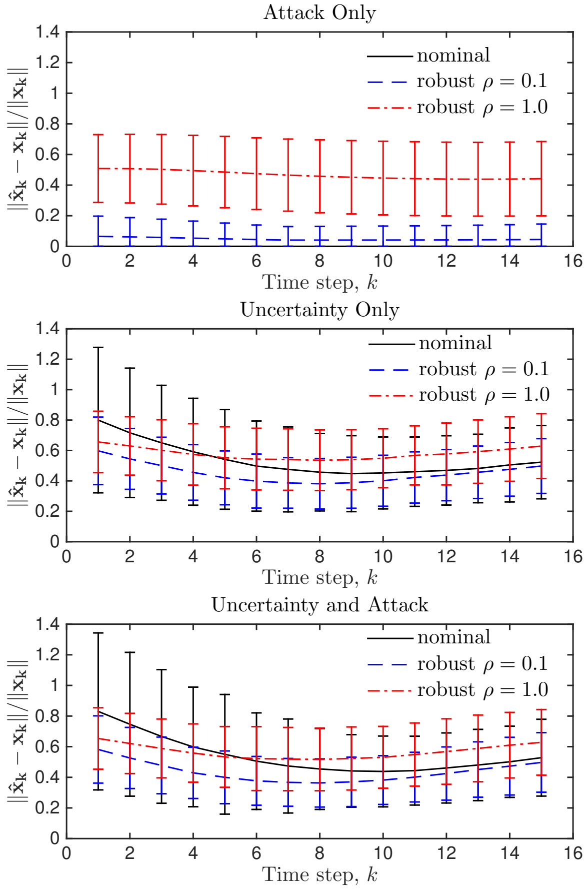


Fig. 5. Normalized errors of the nominal ℓ_1/ℓ_∞ estimator and the robust and resilient ℓ_1/ℓ_∞ estimator (with $\ell_q = \ell_1$, $\lambda = 0.2$ and different values for ρ) under different scenarios for the IEEE 14-bus system. The curves represent mean values and the error bars represent standard deviations.

5) Choices of ℓ_q and ℓ_r in Proposition 1: To find the best values of ℓ_q and ℓ_r in Proposition 1, we fixed the intensity of the modeling errors and the attack signal variance, as well as chose $\lambda = 0.2$ and $\rho = 0.1$. Then, we ran 100 simulations with different sets of attacked sensors K of cardinality $q_s = 3$ and different sets of attacked actuators L of cardinality $q_a = 1$ and with different initial conditions x_0 and perturbations. In particular, we consider different cases for the initial state: (i) with i.i.d. zero-mean unit-variance Gaussian entries (nominal), (ii) with increased variance (increased magnitude), or (iii) with a non-zero mean (offset).

Table I shows the estimate errors of x_0 averaged over 100 simulations for these difference cases of x_0 and for each choice of ℓ_q and ℓ_r from either ℓ_1, ℓ_2 or ℓ_∞. We observe that the mean estimate error is, with the exception of the case with an offset, smaller for $\ell_q = \ell_1$, which coincidentally

TABLE I

MEAN ERROR (VARIED x_0: NOMINAL, INCREASED MAGNITUDE, OFFSET)

	$\ell_r = \ell_1$	$\ell_r = \ell_2$	$\ell_r = \ell_\infty$
$\ell_q = \ell_1$	0.314, 0.250, 0.283	0.296, 0.232, 0.281	0.332, 0.226, 0.364
$\ell_q = \ell_2$	0.330, 0.256, 0.287	0.304, 0.243, 0.275	0.318, 0.236, 0.277
$\ell_q = \ell_\infty$	0.333, 0.256, 0.291	0.312, 0.246, 0.279	0.322, 0.240, 0.283

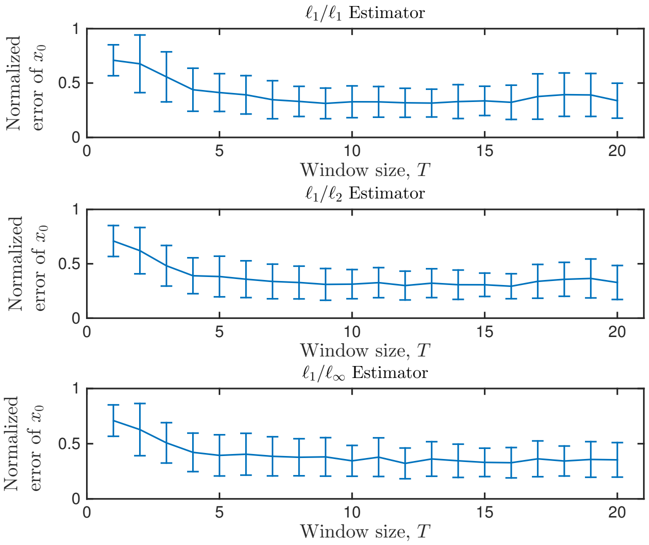


Fig. 6. Performance of the ℓ_1/ℓ_1, ℓ_1/ℓ_2 and ℓ_1/ℓ_∞ estimators (with $\ell_q = \ell_1$, $\lambda = 0.2$, $\rho = 0.1$) on the IEEE 14-bus power network with varying window size T; error bars depict standard deviations of the errors.

also provides the simplification of the optimization problem in Proposition 1 to the one in Corollary 1. And for all the cases with $\ell_q = \ell_1$, $\ell_r = \ell_2$ appears to be the best choice.

6) Varying Window Size T: Lastly, we consider the effects of window size T, i.e., the number of steps/observations that are used for the optimization problem in Proposition 1. We fixed the intensity of the modeling errors and the attack signal variance, as well as chose $\lambda = 0.2$, $\rho = 0.1$ and $\ell_q = \ell_1$. For each value of T, we ran 100 simulations with different sets of attacked sensors K of cardinality $q_s = 3$ and different sets of attacked actuators L of cardinality $q_a = 1$ and with different initial conditions x_0 and perturbations. From Figure 6, we see that the benefit of increasing the window size is high initially but little is gained by increasing the window size beyond approximately $T = 9$, which is less than the number of states $n = 10$. A similar behavior is observed with the nominal estimator for known systems in [9].

VI. Conclusion

We proposed a novel state estimation algorithm that is both resilient to adversarial attacks as well as robust to multiplicative and additive modeling uncertainties/errors. Our approach leverages the equivalence of robustification and regularization in linear regression by constructing suitable uncertainty sets that lead to a tractable optimization solution, such that off-the-shelf convex optimization tools can be readily applied. Moreover, we obtained a novel robust filtering algorithm (as a corollary) when there is no adversarial attack without assuming any priors. We also illustrated the use of cross-validation for determining the hyperparameters of the optimization problem. Using simulations of random systems and an IEEE 14-bus power system, we observed that our robust and resilient estimation approach is effective for decreasing the state estimation errors.

Acknowledgments

The work presented in this paper was supported by the National Science Foundation, grant #1239182.

References

[1] A. A. Cárdenas, S. Amin, and S. Sastry. Research challenges for the security of control systems. In *Proceedings of the 3rd Conference on Hot Topics in Security*, HOTSEC'08, pages 6:1–6:6, 2008.

[2] G. Richards. Hackers vs slackers. *Engineering Technology*, 3(19):40–43, November 2008.

[3] J. P. Farwell and R. Rohozinski. Stuxnet and the future of cyber war. *Survival*, 53(1):23–40, 2011.

[4] M. Zhu and S. Martínez. On distributed constrained formation control in operator-vehicle adversarial networks. *Automatica*, 49(12):3571–3582, 2013.

[5] A. A. Cardenas, S. Amin, and S. Sastry. Secure control: Towards survivable cyber-physical systems. In *International Conference on Distributed Computing Systems Workshops*, pages 495–500, 2008.

[6] Y. Mo and B. Sinopoli. False data injection attacks in control systems. In *Workshop on Secure Control Systems*, 2010.

[7] J. Weimer, S. Kar, and K. H. Johansson. Distributed detection and isolation of topology attacks in power networks. In *Proceedings of the 1st International Conference on High Confidence Networked Systems*, HiCoNS '12, pages 65–72, New York, NY, USA, 2012. ACM.

[8] F. Pasqualetti, F. Dörfler, and F. Bullo. Attack detection and identification in cyber-physical systems. *IEEE Transactions on Automatic Control*, 58(11):2715–2729, Nov 2013.

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

[10] M. Pajic, J. Weimer, N. Bezzo, P. Tabuada, O. Sokolsky, I. Lee, and G. Pappas. Robustness of attack-resilient state estimators. In *ACM/IEEE International Conference on Cyber-Physical Systems (ICCPS)*, pages 163–174, April 2014.

[11] S. Z. Yong, M. Zhu, and E. Frazzoli. Resilient state estimation against switching attacks on stochastic cyber-physical systems. *Conference on Decision and Control (CDC)*, 2015. To appear.

[12] S. Z. Yong, M. Zhu, and E. Frazzoli. Generalized innovation and inference algorithms for hidden mode switched linear stochastic systems with unknown inputs. In *IEEE Conference on Decision and Control*, pages 3388–3394, 2014.

[13] D. Simon. *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. Wiley-Interscience, 1st edition, August 2006.

[14] L. Xie, Y. C. Soh, and C. E. de Souza. Robust Kalman filtering for uncertain discrete-time systems. *IEEE Transactions on Automatic Control*, 39(6):1310–1314, 1994.

[15] G. Calafiore and L. El Ghaoui. Ellipsoidal bounds for uncertain linear equations and dynamical systems. *Automatica*, 40:773–787, 2004.

[16] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust optimization*. Princeton University Press, 2009.

[17] D. Bertsimas, D. B. Brown, and C. Caramanis. Theory and applications of robust optimization. *SIAM review*, 53(3):464–501, 2011.

[18] A. G. Fertis. *A robust optimization approach to statistical estimation problems*. PhD thesis, Massachusetts Institute of Technology, 2009.

[19] D. Bertsimas and M. S. Copenhaver. Characterization of the equivalence of robustification and regularization in linear, median, and matrix regression. *arXiv preprint arXiv:1411.6160*, 2014.

[20] R. Christie. Power Systems Test Case Archive, University of Washington, Electrical Engineering. Online: http://www.ee.washington.edu/research/pstca/, 2000. URL: `http://www.ee.washington.edu/research/pstca/`.

[21] Y. C. Eldar and H. Bölcskei. Block-sparsity: Coherence and efficient recovery. *CoRR*, abs/0812.0329, 2008. URL: `http://arxiv.org/abs/0812.0329`.

[22] H. Xu, C. Caramanis, and S. Mannor. Robust regression and Lasso. In *Advances in Neural Information Processing Systems*, pages 1801–1808, 2009.

[23] F. Pasqualetti, F. Dörfler, and F. Bullo. Cyber-physical attacks in power networks: Models, fundamental limitations and monitor design. In *CDC-ECE*, pages 2195–2201. IEEE, 2011.

[24] F. Pasqualetti, A. Bicchi, and F. Bullo. A graph-theoretical characterization of power network vulnerabilities. In *American Control Conference*, pages 3918–3923. IEEE, 2011.

[25] M. Grant and S. Boyd. Graph implementations for nonsmooth convex programs. In V. Blondel, S. Boyd, and H. Kimura, editors, *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, pages 95–110. Springer-Verlag Limited, 2008.

[26] M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, version 2.1. http://cvxr.com/cvx, March 2014.