2017 IEEE 56th Annual Conference on Decision and Control (CDC)
December 12-15, 2017, Melbourne, Australia

Worst-case analysis of joint attack detection and resilient state estimation

Nicola Forti Giorgio Battistelli Luigi Chisci Bruno Sinopoli

Abstract— **This work investigates the effects of signal attacks possibly combined with network deception attacks injecting fake measurements on stochastic cyber-physical systems. The goal of the attacker is to maximize the estimation error based on the information available about the system and the measurement models, preferably without being detected. This problem is formulated following a worst-case approach characterizing the maximum degradation the attacker can induce at each time instant when a Bayesian filter developed within the random finite set (RFS) framework is employed for simultaneous attack detection and resilient state estimation. A novel concept of error which captures the switching (Bernoulli) nature of the signal attack is proposed as an appropriate distance measure for joint detection–estimation. Furthermore, the notion of stealthiness is introduced in order to derive attack policies useful to synthesize undetectable perturbations that can deceive a Maximum A-posteriori Probability (MAP) detector implemented for security.**

Index Terms— **Cyber-physical systems; integrity attacks; Bayesian state estimation; stealthy attacks.**

I. Introduction

The security of cyber-physical systems (CPSs) is nowadays a topic of paramount importance. In fact, many modern systems for, e.g., electric power generation and distribution, transportation and mobility, building and environmental monitoring/control, health care, and industrial process control, are characterized by a tight interaction of physical and computing processes, interconnected through a communication network, and can therefore be easily compromised by cyber-attackers. For this reason, the design of secure CPSs is attracting great attention [1]–[7]. In particular, the focus of this paper is on secure state estimation, whose aim is to reconstruct the state of the CPS of interest even when it is subject to cyber-physical attacks. Recent work [8] has formulated and solved the problem of detecting a switching signal attack and securely estimating the state of the CPS also in presence of fake measurement injection, by following a Bayesian random set approach. The random set paradigm has been used to model the switching nature of the signal attack and the injection of counterfeit measurements via Bernoulli and, respectively, Poisson random sets. Further, the stochastic Bayesian framework allows to account for several sources of randomness in the estimation (e.g. process and measurement noises, attack signal, counterfeit measurements) in a probabilistic way unlike deterministic attack monitors based on residual analysis. On the other hand, the derivation of a Bayesian filter for joint attack detection and resilient state estimation requires statistical assumptions on the time-correlations and probability distributions of the involved stochastic signals which, especially for the attack signal and fake measurements chosen by cyber-attackers, are very unlikely to hold in practice. In this respect, this paper performs a worst-case analysis of the Bayesian joint attack detector & state estimator presented in [8] and extended in [9] and [10] to distributed settings and, respectively, multiple attack modes. Our intention is to show the inherent robustness of the proposed Bayesian random set filter with respect to mismatches between the simplifying hypothesized modelling assumptions (i.e. whiteness and Gaussianity) on the attack signal and a worst-case attack signal suitably on-line synthesized by the cyber-attacker so as to remain stealthy while maximizing the estimation error. To be more specific, in the considered worst-case analysis, it is assumed that the attacker has perfect knowledge of the states of both the system and the estimator, and also knows the algorithms being used by the CPS monitor for attack detection and state estimation. This is certainly an optimistic situation from the point of view of the attacker (and pessimistic from the CPS monitor viewpoint) and, hence, represents a valid testbed for the effectiveness of the proposed Bayesian approach to secure state estimation. Stealthiness conditions for a Maximum A posteriori Probability (MAP) attack detector will be analyzed and a suitable performance loss to be maximized by the attacker (and minimized by the CPS monitor) will be defined. The main result of the worst-case analysis will be to show that the cyber-attacker, in order to remain stealthy, cannot degrade the CPS monitor performance loss beyond a certain extent. It is worth to point out that previous studies, e.g. [11], [12], have analyzed how the control/estimation performance can be degraded by integrity attacks on CPSs and have characterized the notion of attack stealthiness. Differently from previous work, the present paper aims to provide an analysis of the robustness of CPSs by following a probabilistic approach.

The rest of the paper is organized as follows. Section II deals with the problem setup (system and attack models, background on random sets). Section III reviews the random set Bayesian approach to joint attack detection and state esti-

This work was supported in part by the Department of Energy under Award Number DE-OE0000779 and in part by Ente Cassa di Risparmio di Firenze under Grant 2015-0772.

N. Forti, G. Battistelli, and L. Chisci (nicola.forti,giorgio.battistelli,luigi.chisci@unifi.it) are with the Department of Information Engineering, University of Florence, Via Santa Marta 3, 50139 Florence, Italy.

B. Sinopoli (brunos@ece.cmu.edu) is with the Department of Electrical and Computer Engineering, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA.

mation of [8]. Section IV provides a worst-case performance analysis of the Bayesian joint attack detector-state estimator. Then, Section V investigates a numerical case-study. Finally, Section VI ends the paper with concluding remarks.

II. Problem setup

A. System and attack model

The discrete-time system of interest is

$$x_{k+1} = \begin{cases} f_k^0(x_k) + w_k, & \text{under no attack} \\ f_k^1(x_k, a_k) + w_k, & \text{under attack} \end{cases} \quad (1)$$

where: k is the time index; $x_k \in \mathbb{R}^n$ is the state vector to be estimated; $a_k \in \mathbb{R}^p$, called attack vector, is an unknown input affecting the system under attack; $f_k^0(\cdot)$ and $f_k^1(\cdot,\cdot)$ are known state transition functions that describe the system evolution in the *no attack* and, respectively, *attack* cases; w_k is a random process disturbance also affecting the system. For monitoring purposes, the state of the above system is observed through the measurement

$$y_k = \begin{cases} h_k^0(x_k) + v_k, & \text{under no attack} \\ h_k^1(x_k, a_k) + v_k, & \text{under attack} \end{cases} \quad (2)$$

where: $h_k^0(\cdot)$ and $h_k^1(\cdot,\cdot)$ are known measurement functions that refer to the *no attack* and, respectively, *attack* cases; v_k is a random measurement noise. It is assumed that the measurement y_k is actually delivered to the system monitor with probability $p_d \in (0,1]$, where the non-unit probability is due to several possible reasons like, e.g., temporary denial of service, packet loss, sensor inability to detect or sense the system. The attack modeled in (1)-(2) via the attack vector a_k is usually referred to as *signal* attack. Besides the signal attack, the proposed threat model takes into account the possible presence of malicious *extra packet injections*, already considered in [8] and [13]. This means that, in addition to the system-originated measurement y_k in (2), the system monitor might receive from some cyber-attacker extra fake measurements indistinguishable from the system-originated one. For the subsequent developments, it is convenient to introduce the *attack set* at time k, $\mathcal{A}_k$, which is either equal to the empty set if the system is not under signal attack at time k or to the singleton $\{a_k\}$ otherwise, i.e.

$$\mathcal{A}_k = \begin{cases} \emptyset, & \text{if the system is not under signal attack} \\ \{a_k\}, & \text{otherwise.} \end{cases}$$

Due to the possible presence of the *extra packet injection* attack, it is also convenient to define the *measurement set* at time k

$$\mathcal{Z}_k = \mathcal{Y}_k \cup \mathcal{F}_k \quad (3)$$

where

$$\mathcal{Y}_k = \begin{cases} \emptyset & \text{with probability } 1 - p_d \\ \{y_k\} & \text{with probability } p_d \end{cases} \quad (4)$$

is the set of system-originated measurements and $\mathcal{F}_k$ the finite set of fake measurements.

The problem of joint attack detection and state estimation amounts to jointly estimating, at each time k, the state x_k and signal attack set $\mathcal{A}_k$ given the set of measurements $\mathcal{Z}^k \triangleq \cup_{i=1}^k \mathcal{Z}_i$ up to time k.

B. Random set estimation

An RFS (*Random Finite Set*) $\mathcal{Z}$ over $\mathbb{Z}$ is a random variable taking values in $\mathcal{F}(\mathbb{Z})$, the collection of all finite subsets of $\mathbb{Z}$. The mathematical background needed for Bayesian random set estimation can be found in [14]; here, only the basic concepts needed for the specific problem at hand will be recalled. The statistics of an RFS $\mathcal{Z}$ is completely characterized by the *set density* $f(\mathcal{Z})$, also called FISST (*FInite Set STatistics*) probability density. In fact, given $f(\mathcal{Z})$, the cardinality *probability mass function* $p(m)$ that $\mathcal{Z}$ have $m \geq 0$ elements and the joint PDFs $f(z_1, z_2, \ldots, z_m|m)$ over $\mathbb{Z}^m$ given that $\mathcal{Z}$ have m elements, are obtained as follows:

$$\begin{aligned} p(m) &= \frac{1}{m!}\int_{\mathbb{Z}^m} f(\{z_1,\ldots,z_m\})dz_1\cdots dz_m \\ f(z_1,\ldots,z_m|m) &= \frac{1}{m!\,p(m)}\, f(\{z_1,\ldots,z_m\}). \end{aligned}$$

In order to measure probability over subsets of $\mathbb{Z}$ or compute expectations of random set variables, Mahler [14] introduced the notion of *set integral* for a generic real-valued function $g(\mathcal{Z})$ of an RFS $\mathcal{Z}$ as

$$\int g(\mathcal{Z})\,\delta\mathcal{Z} = g(\emptyset) + \sum_{m=1}^{\infty} \frac{1}{m!}\int g(\{z_1,\ldots,z_m\})\,dz_1\cdots dz_m \quad (5)$$

Two specific types of RFSs, i.e. Bernoulli and Poisson RFSs, will be considered in this work in order to model the attack set $\mathcal{A}_k$ and the set of fake measurements $\mathcal{F}_k$ at a given time k. In particular, the attack set is modeled as a Bernouli RFS which can be either empty or, with some probability $r \in [0,1]$, a singleton $\{a\}$ distributed over $\mathbb{A}$ according to the PDF $\varrho(\cdot)$. Accordingly, its set density is defined as follows:

$$f(\mathcal{A}) = \begin{cases} 1-r, & \text{if } \mathcal{A} = \emptyset \\ r\cdot\varrho(a), & \text{if } \mathcal{A} = \{a\} \end{cases}. \quad (6)$$

III. Bayesian joint attack detector and state estimator

As stated above, the signal attack input is modeled as a Bernoulli random set $\mathcal{A} \in \mathcal{B}(\mathbb{A})$, where $\mathcal{B}(\mathbb{A}) = \emptyset \cup \mathcal{S}(\mathbb{A})$ is a set of all finite subsets of the attack space $\mathbb{A} \subseteq \mathbb{R}^q$, and $\mathcal{S}$ denotes the set of all singletons (i.e., sets with cardinality 1) $\{a\}$ such that $a \in \mathbb{A}$. Further, the state vector to be estimated takes values in the state space $\mathbb{X} \subseteq \mathbb{R}^n$. Hence, for the purpose of joint attack detection and state estimation, it is convenient to introduce the *Hybrid Bernoulli Random Set* (HBRS) $\mathcal{X} \triangleq (\mathcal{A}, x)$, as a new state variable which incorporates the Bernoulli attack random set $\mathcal{A}$ and the random state vector x, taking values in the hybrid space $\mathcal{B}(\mathbb{A}) \times \mathbb{X}$. A HBRS is fully specified by the (signal attack) probability r of $\mathcal{A}$ being a singleton, the PDF $p^0(x)$ defined on the state space $\mathbb{X}$, and the joint PDF $p^1(a,x)$ defined on

the joint attack input-state space $\mathbb{A} \times \mathbb{X}$, i.e.

$$p(\mathcal{A}, x) = \begin{cases} (1-r)\, p^0(x), & \text{if } \mathcal{A} = \emptyset \\ r \cdot p^1(a, x), & \text{if } \mathcal{A} = \{a\} \end{cases} . \qquad (7)$$

In [8], we proposed a Bayesian filter to recursively solve the problem of joint attack detection and resilient state estimation of stochastic CPSs by processing, at each time instant, the current observation set $\mathcal{Z}$ along with all the available information (about the attack and the state) provided by the a-priori hybrid Bernoulli density $p(\mathcal{A}, x)$.

MAP attack detector

For the forthcoming analysis on stealthiness, the criterion adopted for attack detection has to be specified. In this respect, *Maximum A posteriori Probability* (MAP) attack detection will be considered. In particular, the decision about whether $\hat{\mathcal{A}} \neq \emptyset$ or $\hat{\mathcal{A}} = \emptyset$ (the system is under signal attack or not) is based on the maximum of the *a posteriori* probabilities Prob($\mathcal{A} \neq \emptyset | \mathcal{Z}$) and Prob($\mathcal{A} = \emptyset | \mathcal{Z}$) for the two hypotheses. Since this is a binary hypothesis problem, whose outcome is only depending on a specific observation set $\mathcal{Z}$ in $\mathcal{F}(\mathbb{Z})$, the decision rule can be used to divide the overall observation space $\mathcal{F}(\mathbb{Z})$ into two decision regions, $\mathbb{F}_0$ and $\mathbb{F}_1$. Whenever the measurement set falls in $\mathbb{F}_0$, the MAP detector will choose $\hat{\mathcal{A}} = \emptyset$, whereas if $\mathcal{Z}$ falls in $\mathbb{F}_1$, the MAP detector will establish that $\hat{\mathcal{A}} \neq \emptyset$. The MAP decision rule for a Bayesian attack detector based on the measurement model (3) is detailed below. Let $\ell^0(y|x)$ and $\ell^1(y|a, x)$ denote the likelihood functions associated to the measurement model (2) in the no attack and, respectively, attack cases.

Lemma 1: When $\mathcal{Z} \neq \emptyset$, the MAP detector for measurement model (3) assigns $\hat{\mathcal{A}} = \emptyset$ if and only if $\mathcal{Z} \in \mathbb{F}_0$, where $\mathbb{F}_0 = \mathcal{F}(\mathbb{Z}) \setminus \mathbb{F}_1$, $\mathbb{F}_1 = \{\mathcal{Z} \in \mathcal{F}(\mathbb{Z}) : \frac{\alpha_0(\mathcal{Z})}{\alpha_1(\mathcal{Z})} \leq \frac{r}{1-r}\}$ and

$$\alpha_0(\mathcal{Z}) \stackrel{\triangle}{=} 1 - p_d + p_d \sum_{y \in \mathcal{Z}} \ell^0(y|x)/\nu(y) \qquad (8)$$

$$\alpha_1(\mathcal{Z}) \stackrel{\triangle}{=} 1 - p_d + p_d \sum_{y \in \mathcal{Z}} \ell^1(y|a, x)/\nu(y). \qquad (9)$$

Moreover, when $\mathcal{Z} = \emptyset$, the MAP detector assigns $\hat{\mathcal{A}} = \emptyset$ if and only if $r < 1/2$.

Proof: Let us consider the conditional probabilities

$$p_0(\mathcal{Z}) \stackrel{\triangle}{=} \text{Prob}(\mathcal{A} = \emptyset | \mathcal{Z}) \qquad (10)$$

$$p_1(\mathcal{Z}) \stackrel{\triangle}{=} \text{Prob}(\mathcal{A} \neq \emptyset | \mathcal{Z}). \qquad (11)$$

By exploiting the Bayes rule, we can rewrite

$$p_0(\mathcal{Z}) = \text{Prob}(\mathcal{A} = \emptyset)\, \text{Prob}(\mathcal{Z} | \mathcal{A} = \emptyset)/c \qquad (12)$$

$$p_1(\mathcal{Z}) = \text{Prob}(\mathcal{A} \neq \emptyset)\, \text{Prob}(\mathcal{Z} | \mathcal{A} \neq \emptyset)/c \qquad (13)$$

where c is a normalizing factor. For the measurement model (3), the above probabilities take the form (see [8])

$$p_0(\mathcal{Z}) = (1-r)\, e^{-\xi} \prod_{y \in \mathcal{Z}} \nu(y) \left[1 - p_d + p_d \sum_{y \in \mathcal{Z}} \frac{\ell^0(y|x)}{\nu(y)}\right] /c$$

$$p_1(\mathcal{Z}) = r\, e^{-\xi} \prod_{y \in \mathcal{Z}} \nu(y) \left[1 - p_d + p_d \sum_{y \in \mathcal{Z}} \frac{\ell^1(y|a, x)}{\nu(y)}\right] /c$$

Given the measurement set $\mathcal{Z}$, the MAP detector assigns $\hat{\mathcal{A}} = \emptyset$ if $p_0(\mathcal{Z}) > p_1(\mathcal{Z})$, i.e.

$$\frac{\alpha_0(\mathcal{Z})}{\alpha_1(\mathcal{Z})} > \frac{r}{1-r} \qquad (14)$$

where (8)-(9) have been used. Thus, if the detector receives a non-empty set $\mathcal{Z}$, it assigns $\hat{\mathcal{A}} = \emptyset$ only if $\mathcal{Z} \in \mathbb{F}_0$, which is the region in $\mathcal{F}(\mathbb{Z})$ associated to the absence of the attack when $\mathcal{Z}$ is delivered. In the case $\mathcal{Z} = \emptyset$, the MAP detector receives no information through observation of the state, and hence, as it can be easily derived from (14), it will assign $\hat{\mathcal{A}} = \emptyset$ if and only if $r < 1/2$, where r is the a priori probability available on the existence of the attack before the MAP test is carried out.

■

Note that, even though the above MAP test is Bayes-optimal, it may not achieve the minimum mean square error. In addition, as it will be shown in Section IV, the MAP assumption is useful to characterize the worst-case signal attack that achieves maximum error on state and attack estimation, while it is clearly unnecessary when $\mathcal{Z} = \emptyset$. It is also worth pointing out that in the case of no extra packet injection and $p_d = 1$, i.e. when $\mathcal{Z} = \{y\}$ in (3), the MAP criterion (14) leads to the standard likelihood ratio test [15]

$$\frac{\ell^0(y|x)}{\ell^1(y|a, x)} \lessgtr \frac{r}{1-r}. \qquad (15)$$

IV. Worst-case performance analysis

In this section the aim is to find the maximal performance degradation that an attacker can induce on the system through a worst-case static analysis, in which the dynamic model is not taken into account. At each time instant, based on the information available described in Section III, the goal of the attacker is to maximize the negative effects the injected signal attack a can cause, given the state x of the system and the measurement model

$$\mathcal{Z} = \mathcal{Y} \cup \mathcal{F} \qquad (16)$$

with system-originated measurement y given by

$$y = \begin{cases} h^0(x) + v, & \text{under no attack} \\ h^1(x, a) + v, & \text{under attack} \end{cases} . \qquad (17)$$

As further described below, the worst-case analysis can be either addressed as a maximization problem of a suitably defined performance error where the attacker has no constraints on the choice of the attack input, or as a constrained problem in which, in order to avoid detection, the signal attack must fulfill a certain condition to remain *stealthy*.

Although we restrict our attention to the static estimation problem, the proposed analysis gives meaningful insights that can be helpful also in the dynamic case, when the attacker maximizes at each time instant the effects of the signal attack injected on the system through a greedy strategy.

A. Performance loss

We are interested in finding the worst-case performance degradation in terms of the average error between the true random set $\mathcal{X}$ and its estimate $\hat{\mathcal{X}} \triangleq (\hat{\mathcal{A}}, \hat{x})$. In particular, the error is averaged over the measurement set $\mathcal{Z}$ since the attacker is assumed to have no knowledge on $\mathcal{Z}$ when the signal attack is synthesized. Similarly to [16] for random sets, the standard concept of Euclidean error between random vectors can be extended in order to define a generalized metric $e(\mathcal{X}, \hat{\mathcal{X}})$ between two hybrid Bernoulli random sets, which accounts for the possibility of the attack set being empty. In particular, the error on state estimation $e_x(\mathcal{X}, \hat{\mathcal{X}})$ is defined as:

$$e_x\left[(\mathcal{A}, x), (\emptyset, \hat{x})\right] = ||x - \hat{x}^0|| \quad (18)$$

$$e_x\left[(\mathcal{A}, x), (\hat{a}, \hat{x})\right] = ||x - \hat{x}^1|| \quad (19)$$

while the definition of the error $e_a(\mathcal{A}, \hat{\mathcal{A}})$ on the joint detection-estimation of the attack random set is

$$e_a(\emptyset, \emptyset) = 0 \quad (20)$$

$$e_a(a, \hat{a}) = ||a - \hat{a}|| \quad (21)$$

$$e_a(\emptyset, \hat{a}) \stackrel{\triangle}{=} e_{0a} \quad (22)$$

$$e_a(a, \emptyset) \stackrel{\triangle}{=} e_{1a} \quad (23)$$

It can be noticed that (22)-(23) define the error for the two possible cardinality mismatches between the *true* and the estimated attack sets. Different metrics can be used for these errors, such as the OSPA (Optimal SubPattern Assignment) distance [16] which assigns $e_{0a} = e_{1a}$ with $e_{1a} \geq ||a - \hat{a}||$, $\forall a, \hat{a} \in \mathbb{A}$. According to the aforementioned metric, the following mean square error averaged on the measurement set can be defined

$$\sigma^2(\mathcal{A}, x) = \int p(\mathcal{Z}|\mathcal{A}, x) e^2(\mathcal{X}, \hat{\mathcal{X}}(\mathcal{Z}))\, \delta\mathcal{Z} \quad (24)$$

where $\hat{\mathcal{X}}$ is clearly a function of $\mathcal{Z}$, denoted as $\hat{\mathcal{X}}(\mathcal{Z})$, and $p(\mathcal{Z}|\mathcal{A}, x)$ is the likelihood function. The overall error in (24), which accounts for both joint attack detection-estimation and state estimation discrepancies, can be written as

$$\sigma^2(\mathcal{A}, x) = \beta\, \sigma_x^2(\mathcal{A}, x) + (1 - \beta)\, \sigma_a^2(\mathcal{A}, x), \quad \beta \in [0, 1] \quad (25)$$

so as to differently penalize, through the choice of coefficient β, the two sources of error. In particular, the mean square error on state estimation takes the form

$$\sigma_x^2(\mathcal{A}, x) = \begin{cases} \sigma_x^2(\emptyset, x), & \text{if } \mathcal{A} = \emptyset \\ \sigma_x^2(a, x), & \text{if } \mathcal{A} = \{a\} \end{cases}. \quad (26)$$

Next, we derive the performance loss for two measurement models corresponding to the presence of only a signal attack on the system and of the combined action of signal attack and extra packet injection.

No extra packet injection ($\mathcal{Z} = \mathcal{Y}$): First of all, let us consider the mean square error associated to the presence of the signal attack

$$\sigma_x^2(a, x) = \int p(\mathcal{Z}|\{a\}, x)\, e_x^2(a, x, \mathcal{Z})\, \delta\mathcal{Z} \quad (27)$$

in the case of measurement set $\mathcal{Z} = \mathcal{Y}$ (no extra packet injection attack), where $e_x^2(a, x, \mathcal{Z}) \stackrel{\triangle}{=} e_x^2\left[(\mathcal{A} = \{a\}, x), \hat{\mathcal{X}}(\mathcal{Z})\right]$ is the error on state estimation when $\mathcal{A}$ is non-empty. Since $\hat{\mathcal{X}}$ is a function of the measurement set $\mathcal{Z}$, using the definition of integral (5) in (27) leads to

$$\begin{aligned} \sigma_x^2(a, x) &= p(\mathcal{Z} = \emptyset|\{a\}, x)\, e_x^2(a, x, \mathcal{Z} = \emptyset) \\ &\quad + \int p(\mathcal{Z} = \{y\}|\{a\}, x)\, e_x^2(a, x, \mathcal{Z} = \{y\})\, dy \end{aligned} \quad (28)$$

where the likelihood function for $\mathcal{Z} = \mathcal{Y}$ is given by

$$p(\mathcal{Z}|\{a\}, x) = \begin{cases} 1 - p_d, & \text{if } \mathcal{Z} = \emptyset \\ p_d\, \ell^1(y|a, x), & \text{if } \mathcal{Z} = \{y\} \end{cases} \quad (29)$$

$\ell^1(y|a, x)$ being the conventional likelihood of y due to the system under attack a, in state x. Hence, by substituting (29) into (28) we obtain

$$\begin{aligned} \sigma_x^2(a, x) &= (1 - p_d)\, e_x^2(a, x, \mathcal{Z} = \emptyset) \\ &\quad + p_d \int \ell^1(y|a, x)\, e_x^2(a, x, \mathcal{Z} = \{y\})\, dy. \end{aligned} \quad (30)$$

According to Lemma 1, the integral over $\mathbb{Z}$ in (30) can be broken down into two integrations over the distinct regions $\mathbb{Z}_0$ and $\mathbb{Z}_1$. These are the decision regions defined in Lemma 1, restricted to the case $|\mathcal{Z}| = 1$, i.e. $\mathbb{Z}_0 = \mathbb{F}_0 \cap \mathbb{Z}$ and $\mathbb{Z}_1 = \mathbb{F}_1 \cap \mathbb{Z}$. This also allows us to explicitly write the error in (30) on state estimation when $\mathcal{Z} \neq \emptyset$ as

$$e_x^2(a, x, \mathcal{Z} = \{y\}) = \begin{cases} ||x - \hat{x}^0(y)||^2, & \text{if } y \in \mathbb{Z}_0 \\ ||x - \hat{x}^1(y)||^2, & \text{if } y \in \mathbb{Z}_1 \end{cases} \quad (31)$$

while for $\mathcal{Z} = \emptyset$ we have

$$e_x^2(a, x, \mathcal{Z} = \emptyset) = \begin{cases} ||x - \hat{x}^0(y)||^2, & \text{if } r < 1/2 \\ ||x - \hat{x}^1(y)||^2, & \text{otherwise} \end{cases}. \quad (32)$$

After breaking the integral, (30) can be rewritten as

$$\begin{aligned} \sigma_x^2(a, x) &= (1 - p_d)\, e_x^2(a, x, \mathcal{Z} = \emptyset) \\ &\quad + p_d \int_{\mathbb{Z}_0} \ell^1(y|a, x)\, ||x - \hat{x}^0(y)||^2\, dy \\ &\quad + p_d \int_{\mathbb{Z}_1} \ell^1(y|a, x)\, ||x - \hat{x}^1(y)||^2\, dy. \end{aligned} \quad (33)$$

Furthermore, following the same rationale used for the error on state estimation, the mean square error on joint attack detection-estimation in (25) can be written for $\mathcal{A} = \{a\}$ as

$$\begin{aligned} \sigma_a^2(a, x) &= (1 - p_d)\, e_a^2(a, x, \mathcal{Z} = \emptyset) \\ &\quad + p_d \int \ell^1(y|a, x)\, e_a^2(a, x, \mathcal{Z} = \{y\})\, dy. \end{aligned} \quad (34)$$

From Lemma 1, the error for $\mathcal{Z}=\{y\}$ takes the form

$$e_a^2(a,x,\mathcal{Z}=\{y\}) = \begin{cases} e_{1a}^2, & \text{if } y\in\mathbb{Z}_0 \\ ||a-\hat{a}(y)||^2, & \text{if } y\in\mathbb{Z}_1 \end{cases} \quad (35)$$

so that (34) becomes

$$\begin{aligned}\sigma_a^2(a,x) &= (1-p_d)\, e_a^2(a,x,\mathcal{Z}=\emptyset) \quad (36)\\ &+p_d\, e_{1a}^2 \int_{\mathbb{Z}_0} \ell^1(y|a,x)\, dy \\ &+p_d \int_{\mathbb{Z}_1} \ell^1(y|a,x)\, ||a-\hat{a}(y)||^2\, dy\end{aligned}$$

where for $\mathcal{Z}=\emptyset$ we defined

$$e_a^2(a,x,\mathcal{Z}=\emptyset) = \begin{cases} e_{1a}^2, & \text{if } r<1/2 \\ ||a-\hat{a}(y)||^2, & \text{otherwise} \end{cases}. \quad (37)$$

Extra packet injection ($\mathcal{Z}=\mathcal{Y}\cup\mathcal{F}$): In the presence of possible extra fake measurements injection in the communication channel, we consider the observation set (3). In this case, $\mathcal{Y}$ is a Bernoulli random set which depends on whether the system-originated measurement y is delivered or not, while $\mathcal{F}$ models fake measurements as a Poisson random set such that the number of extra packets is Poisson-distributed (see [8] for further details on such random set modeling of fake measurements). For this measurement model, $p(\mathcal{Z}|\mathcal{A},x)$ for a non-empty attack set $\mathcal{A}$ takes the form

$$p(\mathcal{Z}|a,x) = \gamma(\mathcal{F})\left[1-p_d+p_d\sum_{y\in\mathcal{Z}}\frac{\ell^1(y|a,x)}{\nu(y)}\right] \quad (38)$$

where $\gamma(\mathcal{F})=e^{-\xi}\prod_{y\in\mathcal{Z}}\nu(y)$ is the FISST PDF of fake-only measurements with average number ξ and intensity $\nu(\cdot)=\xi\,\kappa(\cdot)$, $\kappa(\cdot)$ being the PDF of the elements in $\mathcal{F}$. By applying the definition of set integral (5) to (27) and substituting (38), we can write

$$\begin{aligned}\sigma_x^2(a,x) &= (1-p_d)e^{-\xi}\, e_x^2(a,x,\mathcal{Z}=\emptyset) \quad (39)\\ &+p_d\sum_{m=1}^{\infty}\frac{e^{-\xi}}{m!}\int_{\mathbb{Z}^m}\sum_{i=1}^{m}\ell^1(y_i|a,x)\prod_{j\neq i}\nu(y_j)\, e_x^2(a,x,\mathcal{Z}\neq\emptyset)dy_{1:m}\\ &+(1-p_d)\sum_{m=1}^{\infty}\frac{e^{-\xi}}{m!}\int_{\mathbb{Z}^m}\prod_{i=1}^{m}\nu(y_i)\, e_x^2(a,x,\mathcal{Z}\neq\emptyset)dy_{1:m}\end{aligned}$$

where we used $p(\emptyset|a,x)=(1-p_d)e^{-\xi}$ for the case $\mathcal{Z}=\emptyset$. Note that the second term on the RHS of (39) represents the event of receiving a set of $|\mathcal{Z}|=m$ packets $y_{1:m}=\{y_1,\ldots,y_m\}$, $m=1,2,\ldots$ which include the system-originated measurement together with a set of fake observations. In this case, $p(\mathcal{Z}|a,x)$ with $\mathcal{Y}=\{y\}$ is the union of disjoint events that each point of $\mathcal{Z}$ is system-originated and the rest are fake measurements, i.e. $p_d\sum_{y\in\mathcal{Z}}\ell^1(y|a,x)e^{-\xi}\prod_{z\in\mathcal{Z}\setminus y}\nu(z)$. Finally, the third term on the RHS models the arrival of only fake packets, i.e. the event $\mathcal{Z}=\mathcal{F}$. Next, let us define

$$\begin{aligned}\ell_m^1(\mathcal{Z}|a,x) &\triangleq \quad (40)\\ &\frac{1}{m\,\xi^{m-1}}\sum_{m=1}^{\infty}\frac{e^{-\xi}}{m!}\int_{\mathbb{Z}^m}\sum_{i=1}^{m}\ell^1(y_i|a,x)\prod_{j\neq i}\nu(y_j)dy_{1:m}\end{aligned}$$

the probability density of $\mathcal{Z}$ conditioned on $|\mathcal{Z}|=m$ and on the fact that one of the points of $\mathcal{Z}$ is system-originated. Notice that the normalizing factor in (40) is $m\,\xi^{m-1}=\int\sum_{i=1}^{m}\ell^1(y_i|a,x)\prod_{j\neq i}\nu(y_i)dy_{1:m}$. Then, (39) can be rewritten as

$$\begin{aligned}\sigma_x^2(a,x) &= (1-p_d)e^{-\xi}\, e_x^2(a,x,\mathcal{Z}=\emptyset) \quad (41)\\ &+p_d\sum_{m=1}^{\infty}\frac{m\,\xi^{m-1}e^{-\xi}}{m!}\int_{\mathbb{Z}^m}\ell_m^1(\mathcal{Z}|a,x)\, e_x^2(a,x,\mathcal{Z}\neq\emptyset)dy_{1:m}\\ &+(1-p_d)\sum_{m=1}^{\infty}\frac{e^{-\xi}}{m!}\int_{\mathbb{Z}^m}\prod_{i=1}^{m}\nu(y_i)\, e_x^2(a,x,\mathcal{Z}\neq\emptyset)dy_{1:m}.\end{aligned}$$

On the basis of Lemma 1, when a m–point measurement set is received, the MAP detector will assign $\hat{\mathcal{A}}=\emptyset$ iff $\mathcal{Z}\in\mathbb{Z}_0^m$ and $\hat{\mathcal{A}}\neq\emptyset$ iff $\mathcal{Z}\in\mathbb{Z}_1^m$, where $\mathbb{Z}_0^m,\mathbb{Z}_1^m$ are the decision regions defined in Lemma 1, restricted to the case $|\mathcal{Z}|=m$, i.e. $\mathbb{Z}_0^m=\mathbb{F}_0\cap\mathbb{Z}^m$ and $\mathbb{Z}_1^m=\mathbb{F}_1\cap\mathbb{Z}^m$. Hence, the expressions of the errors in (41) can be computed as follows

$$e_x^2(a,x,\mathcal{Z}\neq\emptyset) = \begin{cases} ||x-\hat{x}^0(\mathcal{Z})||^2, & \text{if } \mathcal{Z}\in\mathbb{Z}_0^m \\ ||x-\hat{x}^1(\mathcal{Z})||^2, & \text{if } \mathcal{Z}\in\mathbb{Z}_1^m \end{cases} \quad (42)$$

while for $\mathcal{Z}=\emptyset$ we have

$$e_x^2(a,x,\mathcal{Z}=\emptyset) = \begin{cases} ||x-\hat{x}^0(\mathcal{Z})||^2, & \text{if } r<1/2 \\ ||x-\hat{x}^1(\mathcal{Z})||^2, & \text{otherwise} \end{cases}. \quad (43)$$

Moreover, by dividing the integral in (41) into two integrals over $\mathbb{Z}_0^m$ and $\mathbb{Z}_1^m$, we obtain

$$\begin{aligned}\sigma_x^2(a,x) &= (1-p_d)e^{-\xi}\, e_x^2(a,x,\mathcal{Z}=\emptyset) \quad (44)\\ &+p_d\sum_{m=1}^{\infty}\frac{m\,\xi^{m-1}e^{-\xi}}{m!}\int_{\mathbb{Z}_0^m}\ell_m^1(\mathcal{Z}|a,x)||x-\hat{x}^0(\mathcal{Z})||^2dy_{1:m}\\ &+p_d\sum_{m=1}^{\infty}\frac{m\,\xi^{m-1}e^{-\xi}}{m!}\int_{\mathbb{Z}_1^m}\ell_m^1(\mathcal{Z}|a,x)||x-\hat{x}^1(\mathcal{Z})||^2dy_{1:m}\\ &+(1-p_d)\sum_{m=1}^{\infty}\frac{e^{-\xi}}{m!}\int_{\mathbb{Z}_0^m}\prod_{i=1}^{m}\nu(y_i)||x-\hat{x}^0(\mathcal{Z})||^2dy_{1:m}\\ &+(1-p_d)\sum_{m=1}^{\infty}\frac{e^{-\xi}}{m!}\int_{\mathbb{Z}_1^m}\prod_{i=1}^{m}\nu(y_i)||x-\hat{x}^1(\mathcal{Z})||^2dy_{1:m}.\end{aligned}$$

In an analogous way, the mean square error on joint attack detection-estimation in the case of extra packet injection attack can be written for $\mathcal{A}=\{a\}$ as

$$\begin{aligned}\sigma_a^2(a,x) &= (1-p_d)\, e^{-\xi}\, e_a^2(a,x,\mathcal{Z}=\emptyset) \quad (45)\\ &+p_d\, e_{1a}^2\sum_{m=1}^{\infty}\frac{m\,\xi^{m-1}e^{-\xi}}{m!}\int_{\mathbb{Z}_0^m}\ell_m^1(\mathcal{Z}|a,x)dy_{1:m}\\ &+p_d\sum_{m=1}^{\infty}\frac{m\,\xi^{m-1}e^{-\xi}}{m!}\int_{\mathbb{Z}_1^m}\ell_m^1(\mathcal{Z}|a,x)\,||a-\hat{a}(\mathcal{Z})||^2\,dy_{1:m}\\ &+(1-p_d)\, e_{1a}^2\sum_{m=1}^{\infty}\frac{e^{-\xi}}{m!}\int_{\mathbb{Z}_0^m}\prod_{i=1}^{m}\nu(y_i)dy_{1:m}\\ &+(1-p_d)\sum_{m=1}^{\infty}\frac{e^{-\xi}}{m!}\int_{\mathbb{Z}_1^m}\prod_{i=1}^{m}\nu(y_i)\,||a-\hat{a}(\mathcal{Z})||^2\,dy_{1:m}\end{aligned}$$

where we used, according to Lemma 1, the following error on the attack detection-estimation for $\mathcal{Z} \neq \emptyset$:

$$e_a^2(a,x,\mathcal{Z}\neq\emptyset) = \begin{cases} e_{1a}^2, & \text{if } \mathcal{Z} \in \mathbb{Z}_0^m \\ ||a-\hat{a}(\mathcal{Z})||^2, & \text{if } \mathcal{Z} \in \mathbb{Z}_1^m \end{cases} \quad (46)$$

while the error for $\mathcal{Z} = \emptyset$ in (45) takes the values

$$e_a^2(a,x,\mathcal{Z}=\emptyset) = \begin{cases} e_{1a}^2, & \text{if } r < 1/2 \\ ||a-\hat{a}(\mathcal{Z})||^2, & \text{otherwise} \end{cases} . \quad (47)$$

B. Worst-case performance loss and stealthiness constraint

Suppose now the attacker knows the state x of the system and the measurement model (16), then the worst-case performance loss on Bayesian estimation for $\mathcal{A} \neq \emptyset$ can be found by solving the optimization problem

$$\max_a \quad \sigma^2(a,x) \quad (48)$$

Clearly, unless the signal attack has a specific structure or satisfies some constraints, the worst-case performance can even be infinitely large. The situation is however different when the action of an attacker injecting a signal attack on the system to achieve the maximal performance loss in terms of Bayesian estimation is constrained by the condition on the attack being *stealthy*. In such a case, the goal of the attacker will be to find the worst-case error that can be provoked, without being detected by a decision maker based on the MAP criterion as the one presented in Section III. If on one hand the requirement of being stealthy keeps the monitoring system unaware of the attack presence, on the other it will inevitably narrow down the achievable performance deterioration. Assuming that the attacker knows the MAP decision rule (14), and due to the uncertainty about the measurement set $\mathcal{Z}$, the aim will be to synthesize a signal attack which guarantees a certain probabilistic level of *stealthiness*. This idea can be formalized through the following definition.

Definition 1: **ϵ–stealthiness:** Given $\epsilon \in (0,1)$, a signal attack $\mathcal{A} \neq \emptyset$ is *ϵ–stealthy* if

$$P(a,x) > 1-\epsilon \quad (49)$$

where $P(a,x) = \text{Prob}(\hat{\mathcal{A}} = \emptyset | \mathcal{A} = \{a\}, x)$.

In the presence of extra fake packets, $P(a,x)$ in (49) is given by

$$\begin{aligned} P(a,x) &= p_d \sum_{m=1}^{\infty} \frac{m\,\xi^{m-1}e^{-\xi}}{m!} \int_{\mathbb{Z}_0^m} \ell_m^1(\mathcal{Z}|a,x)\,dy_{1:m} \\ &+(1-p_d)\sum_{m=1}^{\infty}\frac{e^{-\xi}}{m!}\int_{\mathbb{Z}_0^m}\prod_{i=1}^{m}\nu(y_i)\,dy_{1:m} \end{aligned} \quad (50)$$

which reduces to $P(a,x) = p_d \int_{\mathbb{Z}_0} \ell^1(y|a,x)\,dy$ for $\mathcal{Z} = \mathcal{Y}$.

Under the stealthiness constraint (49), the worst-case performance loss on state estimation for $\mathcal{A} \neq \emptyset$ can be obtained as the solution of the problem

$$\begin{aligned} \max_a \quad & \sigma_x^2(a,x) \\ \text{subject to} \quad & P(a,x) > 1-\epsilon \end{aligned} \quad (51)$$

It is worth pointing out that, in solving the constrained problem (51), the attacker must reach a compromise between the dual objectives of worsening the estimation performance and guaranteeing stealthiness of the signal attack with some level of confidence. Note that, in (51) we considered only the error on state estimation, since the intention of deceiving the attack detection is already taken into account by means of constraint (49).

V. Illustrative case-study

In this section we validate the developed analysis for the special class of linear Gaussian SISO models, for which it is possible to derive a closed-form solution to problem (51). In general, no analytical solution is admitted and one can resort to Monte Carlo integration methods [17] which rely on random sampling to numerically compute integrals. Specifically, we consider an observation model with no extra packet injection and $p_d = 1$, i.e.

$$\mathcal{Z} \;=\; \{y\} \quad (52)$$

where y is generated as

$$y = \begin{cases} c\,x + v, & \text{under no attack} \\ c\,x + h\,a + v, & \text{under attack} \end{cases} \quad (53)$$

For such a system, the MAP decision rule (14) reduces to the standard likelihood ratio test (15)

$$\frac{\ell^0\big(y|\hat{x}^0(y)\big)}{\ell^1\big(y|\hat{a}(y),\hat{x}^1(y)\big)} \lessgtr \frac{r}{1-r} \quad (54)$$

where $\hat{x}^0(y), \hat{x}^1(y)$ and $\hat{a}(y)$ are the available a posteriori estimates of the state conditioned on the fact of being under nominal operation, under attack and of the signal attack, respectively. Given x, in the scalar case a turns out to be ϵ–stealthy if $P(a,x) = \int_{\mathbb{Z}_0} \ell^1(y|a,x)\,dy$ satisfies condition (49), where $\mathbb{Z}_0 = [y_{\min}, y_{\max}]$ is a suitable interval in $\mathbb{Z}$. In order to solve the constrained problem (51), the attacker seeks to maximize the performance error

$$\begin{aligned} \sigma_x^2(a,x) \;=\; & \int_{\mathbb{Z}_0} \ell^1(y|a,x)\,(x-\hat{x}^0(y))^2\,dy \\ & + \int_{\mathbb{Z}_1} \ell^1(y|a,x)\,(x-\hat{x}^1(y))^2\,dy. \end{aligned} \quad (55)$$

subject to $\int_{\mathbb{Z}_0} \ell^1(y|a,x)\,dy > 1-\epsilon$. For the simulations we considered the prior distributions $\mathcal{N}(\bar{x}^0, \Sigma_x^0), \mathcal{N}(\bar{x}^1, \Sigma_x^1)$ and $\mathcal{N}(\bar{a}, \Sigma_a)$. In addition, we set $r = 0.2$, $\bar{x}^0 = \bar{x}^1 = 1$, $\bar{a} = 0.5$, $\Sigma_x^0 = \Sigma_x^1 = \Sigma_a = 2$, $c = 1$, $h = 2$, $x = 1.5$ and sensor noise $v \sim \mathcal{N}(0, \sigma_v^2)$ with $\sigma_v^2 = 0.05$. Note that $\hat{x}^0(y), \hat{x}^1(y)$ in (55) are the state estimates provided by a standard static Minimum Mean-Square Estimator (MMSE), and $\hat{a} = (y - c\,\hat{x}^1)/h$. In this practical case-study, all the functions are evaluated by sampling the value a on 2001 evenly spaced points in the interval $[-10, 10]$. Fig. 1 (a) shows the likelihood ratio test (54), upon which the decision of the MAP detector about the presence/absence of the signal attack a is based. As it can be seen in Fig. 1 (b), for a given x, the detector will assign $\hat{\mathcal{A}} =$

$\emptyset$ whenever the measurement $y(a,x)$ is generated with a satisfying $\ell^0(y|\hat{x}^0)/\ell^1(y|\hat{a},\hat{x}^1) > r/1-r$ (red values of y centered in $\bar{y}^0 = c\,\bar{x}^0$). The attacker, with no knowledge of y, will exploit the information about the state and the MAP decision rule to synthesize an ϵ–stealthy attack a such that $P(a,x) = \int_{\mathbb{Z}_0} \ell^1(y|a,x)\,dy > 1-\epsilon$. The resulting ϵ–stealthiness constraint is plotted in Fig. 2 (a) for $\epsilon = 0.1$. We can observe that, due to the abrupt transition between stealthy and unstealthy regions, values of the signal attack too close to the corresponding threshold might generate measurements falling in $\mathbb{Z}_1$, and hence produce the decision $\hat{\mathcal{A}} \neq \emptyset$. On the other hand, the sharp behavior of $P(a,x)$ ensures that for a large interval of values around the midpoint of the ϵ–stealthy region, corresponding to the red circles in Fig. 2 (a), an undetected signal attack will be produced with very high confidence. Finally, Fig. 2 (b) shows the performance loss $\sigma^2_{x,0}(a,x) = \int_{\mathbb{Z}_0} \ell^1(y|a,x)\,||x-\hat{x}^0(y)||^2\,dy$ associated to the attack being undetected, plotted in the region of ϵ–stealthiness.

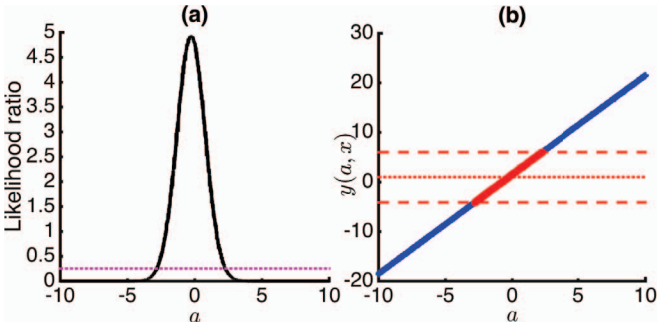

Fig. 1: (a) Likelihood ratio test as a function of a. (b) System-generated measurement y as a function of a.

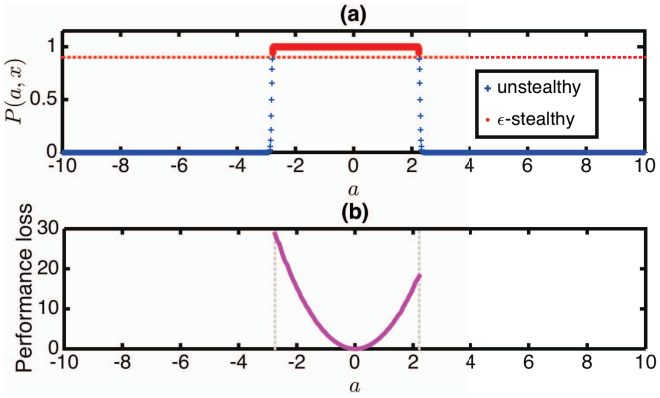

Fig. 2: (a) ϵ–stealthiness constraint as a function of a ($\epsilon = 0.1$). (b) Performance loss $\sigma^2_{x,0}(a,x)$ in the ϵ–stealthiness region.

VI. Conclusions

The paper has addressed important issues concerning the monitoring of a cyber-physical system where, on one hand, a system monitor aims to simultaneously detect the presence of attacks and securely estimate the state of the CPS relying on a Bayesian approach while, on the other hand, an attacker attempts to compromise as much as possible such tasks by injecting a suitably designed attack signal into the CPS. In particular, a performance loss, averaged over the observations, has been introduced in order to measure the joint *attack detection & state estimation* performance of the CPS monitor. Further, a probabilistic notion of stealthiness with a certain degree of confidence has been defined and stealthiness conditions for a MAP detector have been investigated. A worst-case scenario, where the attacker knows both the true system state and the CPS monitor's (state & attack vector) estimates and tries to maximize the performance loss, has been analyzed. The main result of this analysis has been to show that if the attacker wants to remain stealthy with some degree of confidence, then it cannot degrade too much the CPS monitor's performance. This clearly implies that, despite its probabilistic nature, a Bayesian approach to CPS monitoring has an intrinsic degree of robustness.

References

[1] F. Pasqualetti, F. Dörfler, and F. Bullo, "Attack detection and identification in cyber-physical systems," *IEEE Transactions on Automatic Control*, vol. 58, no. 11, pp. 2715–2729, 2013.

[2] H. Fawzi, P. Tabuada, and S. Diggavi, "Secure estimation and control for cyber-physical systems under adversarial attacks," *IEEE Transactions on Automatic Control*, vol. 59, no. 6, pp. 1454–1467, 2014.

[3] Y. Shoukry, A. Puggelli, P. Nuzzo, A. L. Sangiovanni-Vincentelli, S. A. Seshia, and P. Tabuada, "Sound and complete state estimation for linear dynamical systems under sensor attacks using satisfiability modulo theory solving," in *Proc. American Control Conference*, pp. 3818–3823, Chicago, IL, USA, 2015.

[4] A. Teixeira, I. Shames, H. Sandberg, and K. H. Johansson, "A secure control framework for resource-limited adversaries," *Automatica*, vol. 51, no. 1, pp. 135–148, 2015.

[5] M. Pajic, P. Tabuada, I. Lee, and G. J. Pappas, "Attack-resilient state estimation in the presence of noise," in *Proc. 54th IEEE Conference on Decision and Control*, pp. 5827–5832, Osaka, Japan, 2015.

[6] S. Yong, M. Zhu, and E. Frazzoli, "Resilient state estimation against switching attacks on stochastic cyber-physical systems," in *Proc. 54th IEEE Conference on Decision and Control*, pp. 5162–5169, Osaka, Japan, 2015.

[7] Y. Mo and B. Sinopoli, "Secure estimation in the presence of integrity attacks," *IEEE Transactions on Automatic Control*, vol. 60, no. 4, pp. 1145–1151, 2015.

[8] N. Forti, G. Battistelli, L. Chisci, and B. Sinopoli, "A Bayesian approach to joint attack detection and resilient state estimation," in *Proc. 55th IEEE Conference on Decision and Control*, Las Vegas, NV, USA, 2016.

[9] N. Forti, G. Battistelli, L. Chisci, S. Li, B. Wang, and B. Sinopoli, "Distributed joint attack detection and secure state estimation," *IEEE Trans. on Signal and Information Processing over Networks*, 2017.

[10] N. Forti, G. Battistelli, L. Chisci, and B. Sinopoli, "Secure state estimation of cyber-physical systems under switching attacks," in *Proc. 20th IFAC World Congress*, Toulouse, France, 2017.

[11] C. Z. Bai, F. Pasqualetti, and V. Gupta, "Security in stochastic control systems: Fundamental limitations and performance bounds," in *Proc. American Control Conference*, pp. 195–200, Chicago, IL, USA, 2015.

[12] Y. Mo and B. Sinopoli, "On the performance degradation of cyber-physical systems under stealthy integrity attacks," *IEEE Transactions on Automatic Control*, vol. 61, no. 9, pp. 2618–2624, 2015.

[13] Q. Gu, P. Liu, S. Zhu, and C.-H. Chu, "Defending against packet injection attacks in unreliable ad hoc networks," in *Proc. IEEE Global Telecommunications Conference*, vol. 3, pp. 1837–1841, St. Louis, MO, USA, 2005.

[14] R. P. S. Mahler, *Statistical multisource multitarget information fusion*. Artech House, 2007.

[15] H. Van Trees, *Detection, Estimation, and Modulation Theory*. Wiley, 2004.

[16] M. Rezaeian and B. N. Vo, "Error bounds for joint detection and estimation of a single object with random finite set observation," *IEEE Trans. on Signal Processing*, vol. 58, no. 3, pp. 1493–1506, 2010.

[17] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*. Springer-Verlag New York, 2005.