Random Bad State Estimator to Address False Data Injection in Critical Infrastructures

Giulio Masetti*, Silvano Chiaradonna*, Leonardo Robol*†, Felicita Di Giandomenico*
* ISTI-CNR, Pisa, Italy, † University of Pisa, Italy, {giulio.masetti, silvano.chiaradonna, felicita.digiandomenico}@isti.cnr.it
and leonardo.robol@unipi.it

***Abstract*—Given their crucial role for a society and economy, an essential component of critical infrastructures is the Bad State Estimator (BSE), responsible for detecting malfunctions affecting elements of the physical infrastructure. In the past, the BSE has been conceived to mainly cope with accidental faults, under assumptions characterizing their occurrence. However, evolution of the addressed systems category consisting in pervasiveness of ICT-based control towards increasing smartness, paired with the openness of the operational environment, contributed to expose critical infrastructures to intentional attacks, e.g. exploited through False Data Injection (FDI). In the flow of studies focusing on enhancements of the traditional BSE to account for FDI attacks, this paper proposes a new solution that introduces randomness elements in the diagnosis process, to improve detection abilities and mitigate potentially catastrophic common-mode errors. Differently from existing alternatives, the strength of this new technique is that it does not require any additional components or alternative source of information with respect to the classic BSE. Numerical experiments conducted on two IEEE transmission grid tests, taken as representative use cases, show the applicability and benefits of the new solution.**

A. Acronyms and Symbols

BSE	Bad State Estimator
CDF	Cumulative Distribution Function
FDI	False Data Injection
R-BSE	Random Bad State Estimation
SE	State Estimator
$a \in \mathbb{R}^m$	Attack vector
$\alpha \in [0,1]$	Significance level
$B \in \mathbb{R}^{m \times m}$	Change matrix
$c \in \mathbb{R}^n$	State difference vector
$e \in \mathbb{R}^m$	Measurements error vector
$\hat{e} \in \mathbb{R}^m$	Estimation error vector
$E \in \mathbb{R}^{n \times m}$	Estimator matrix
$\hat{e}_a \in \mathbb{R}^m$	False estimation error vector
$\epsilon \geq 0$	Distance from $\mathfrak{X}$, model defender's ignorance
$G \in \mathbb{R}^{n \times n}$	Gain matrix
$h : \mathbb{R}^n \to \mathbb{R}^m$	Physical model
$H \in \mathbb{R}^{m \times n}$	Measurement matrix
$J \in \mathbb{R}$	Weighted sum of squares
$\hat{J} \in \mathbb{R}$	Basic BSE statistic
$\tilde{J} \in \mathbb{R}$	R-BSE statistic
$\tilde{J}_a \in \mathbb{R}$	False R-BSE statistic
$\hat{J}_a \in \mathbb{R}$	False basic BSE statistic
$k \in \mathbb{N}$	Dimension of $\mathfrak{X}$
$m \in \mathbb{N}$	Number of measurements
$M \in \mathbb{R}^{n \times n}$	Confusion matrix
$n \in \mathbb{N}$	Number of state variables
$r \in \mathbb{R}^m$	Residue vector for R-BSE
$r_a \in \mathbb{R}^m$	False residue vector for R-BSE
$\Sigma \in \mathbb{R}^{m \times m}$	Variance matrix
$\hat{\Sigma} \in \mathbb{R}^{m \times m}$	Covariance matrix
$\tau \in \mathbb{R}$	Hypothesis threshold
$\hat{W} \in \mathbb{R}^{m \times m}$	Derived weight matrix
$W \in \mathbb{R}^{m \times m}$	Weight matrix
$x \in \mathbb{R}^n$	State vector
χ^2	Chi-squared distribution, i.e., the norm 2 of a normal vector
$\mathfrak{X} \subsetneq \mathbb{R}^n$	Vector space spanned by x
$\hat{x}_a \in \mathbb{R}^n$	Estimated false state vector
$x_a \in \mathbb{R}^n$	False state vector
$\hat{x} \in \mathbb{R}^n$	Estimated state vector
$z \in \mathbb{R}^m$	Measurements vector
$z_a \in \mathbb{R}^m$	False measurements vector
$\hat{z} \in \mathbb{R}^m$	Estimated measurements vector

I. Introduction

Availability, reliability or resilience of modern cyber-physical systems highly depend on the quality of data provided by the sensors the system is equipped with. When the state of the physical part of the system is not under direct observation by the cyber control, because for instance the system is distributed on a vast territory, measurements are taken by sensors and sent to a control center, where the physical state is inferred. The component that gathers sensors input and processes them, relying on a model that takes into account measurement errors, is called State Estimator (SE) [1]. The outcome of the SE is in turn exploited to govern the system, so the issue of error detection[1] in the data sent from sensors to the control center is crucial, since undetected errors tend to lead to system failures. These errors originate from faults of different kinds and likelihoods, and are addressed differently depending on which system is considered.

This paper targets critical infrastructures, typical cyber-physical systems providing essential services for the functioning

[1]An *error* is defined as the part of the system's total state that may lead to a *failure* [2]. In this paper, when the word *state* is used it is understood *physical state*. Notice that *measurement errors* are determined by physical instruments' sensitivity and are part of the model, so are not errors in the sense of [2]. To disambiguate, in the paper "error(s)" is used with the meaning as in [2], and only errors in data sent from sensors to the control center are considered, as it will be clarified in Section II-B; whereas, "measurements error(s)" have the standard physical meaning and "gross error" indicates the overall SE inaccuracy.

of a society and economy. They are built to operate for decades, undergoing major modifications over a time scale of several years. The error detection mechanisms in place in this sector, that is the detection mechanism that corresponds to SE, is called BSE [1]. The increasing pervasiveness of ICT technologies in all critical infrastructures (electricity, gas, oil, water, etc.) together with a mutated social and legislation context, drastically changed the landscape where those infrastructures operate. Specifically, there has been a transition from a relatively secure operational environment, where faults were mostly accidental (random), almost always involving the (physical) measurement part of a sensor following a characteristic pattern (some transient faults with increasing frequency followed by a permanent fault), to an insecure one, where intentional (deterministic) faults caused by well targeted physical and cyber attacks are becoming increasingly frequent. In particular, common-mode errors, i.e. erroneous but coincident outcomes, affecting employed sensors were quite unlikely in the past (due to geographical displacement of the devices), resulting in a probability of undetected common-mode considered low enough not to deserve special attention (details in Section II-A). Nowadays, instead, FDI, where sensors are intruded and their cyber part altered in such a way that a coordinated attack can lead to undetected common-mode errors (details in Section II-B), is considered a plausible event. The consequences of such an attack can be devastating (e.g., [3]).

In this paper, an evolution of the classic and widely adopted BSE [1] is presented. The classic BSE was not developed to address FDI [4], which motivated the investigation on how to advance it towards including also FDI management, in addition to the detection of errors generated by accidental faults. The logic guiding the development of the new Random Bad State Estimation (R-BSE) is to introduce a minimal number of changes to the classic solution to: i) reduce the cost in terms of changes/additions to the currently adopted configuration; ii) avoid/minimize potential revision of the cyber control system to account for new components and related dependencies; iii) minimize the need for more sophisticated expertise for its management; iv) promote wider applicability of more robust BSE for the reasons just listed at the previous points. The idea at the basis of R-BSE is to introduce random changes in the process to avoid the fully deterministic pattern of BSE, easily exploitable by an attacker to compromise carefully selected measurements in an undetected manner. The randomness introduced acts as an element of confusion for the attacker, who no longer has the certainty on how to intrude on the system by mimicking behavior that cannot be detected. The result is thus a lower probability of undetected common-mode error.

The running example is an electrical transmission grid, but similar reasoning applies to other critical infrastructures. The main reasons for the choice are: • there exist a vast literature on the subject; • the physical model commonly adopted is a linear one (e.g., measurements: power injected at some buses, i.e., nodes of the electrical grid, and power flowing from some buses to others; state: voltage angles at all but one buses). This simplifies the formal treatment, and in any cases prepares the ground to address nonlinear models; • recently FDI attacks to this infrastructure have been reported.

Despite research in this topic has investigated a variety of directions, we believe that R-BSE is simple yet effective enough to be considered an interesting and affordable alternative. Moreover, its simplicity is also expected to favor easy composition with other BSE (see Section V), if synergistic configurations would result appropriate for the specific application at hand.

The rest of the paper is structured as follows. First, Section II summarizes relevant context and recalls the classic BSE, as needed to understand the new technique R-BSE, that is then illustrated in Section III. In particular, Sections II-A and III-A are focused on the analysis of the relevant probability distributions involved in classic BSE and R-BSE, respectively. Section IV presents numerical evaluations that confirm the expected properties of the proposed R-BSE. Section V briefly discusses related work. Section VI draws conclusions and outlines future work.

II. Context

In this section, relevant context necessary to understand at high level the role and limitations of the classic BSE is summarized.

In general terms, BSE is outliers detection [5], and the classic BSE is an univariate statistical approach [6]: starting from the measurements, assumed to have the same distribution and being independent, a real valued random variable that captures SE's *gross error* (see Footnote 1) is defined and tested against a threshold; the probability of the event "gross error greater than the threshold" is so low that this event is considered an anomaly. All the involved distributions are supposed to be known.

Figure 1 sketches both SE and BSE for the linear model so it can guide the reader through the quite complicated notation. Consider a stationary physical system whose state is represented by $x \in \mathfrak{X} \subset \mathbb{R}^n$. Being the system stationary, its state is subject to small changes during long periods of time, so $\mathfrak{X}$ is a *small* vector subspace of $\mathbb{R}^n$ comprising the equilibrium point. Once in a while the system jumps to a different equilibrium point, and then $\mathfrak{X}$ changes accordingly. The precise state x is not directly knowable, even though the control center knows (a good approximation of) $\mathfrak{X}$, thus a set of measurements $z \in \mathbb{R}^m$ are taken (all at the same time) by sensors, sent to the control center and an estimate $\hat{x}$ of the state is inferred from the relation $z = h(x) + e$, where $e \in \mathbb{R}^m$ is the vector of measurement errors (a random vector) and $h : \mathbb{R}^n \to \mathbb{R}^m$ is usually determined exploiting physical *first principles*, and depends on the specific context of application. This is the SE. As an example, consider the DC model of an electrical transmission grid [1]. Here x comprises the voltage angle at every node of the grid except for one of them that is considered as a reference (angle equal to zero), z can comprise nodal powers, power flowing through lines and others, depending on which measurement and ICT components are available (a concrete example is depicted in Figure 2). In this case, the model is linear, i.e., $h(x) = Hx$, where $H \in \mathbb{R}^{m \times n}$.

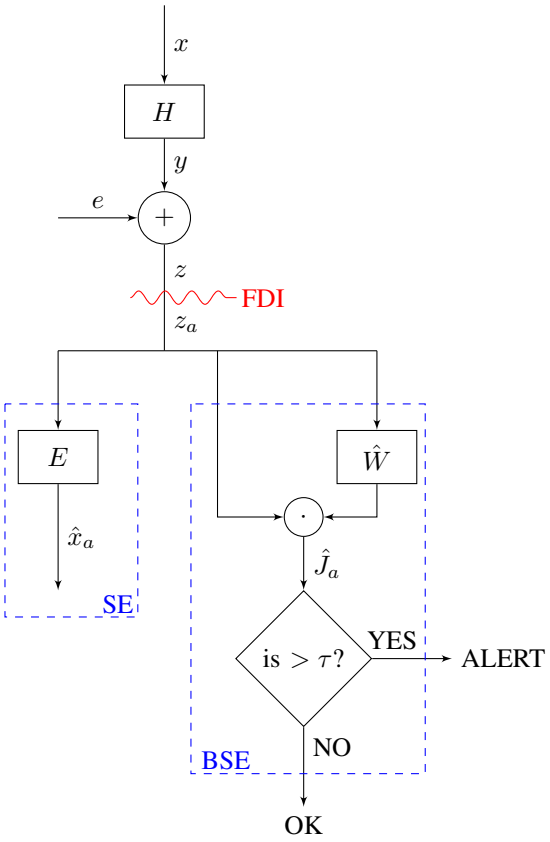


Fig. 1. Pictorial representation of on-line computations performed by SE and classic BSE.

For enhancing both awareness/observability[2] and fault tolerance, m is chosen greater than n. The role of SE is to tolerate unavoidable measurement errors and the presence of sensor faults.

Once the estimate $\hat{x}$ is available, it is possible to compute $\hat{z} := h(\hat{x}) = H\hat{x}$ and then evaluate the estimation error $\hat{e} := \hat{z} - z$. Intuitively, SE and BSE move back and forth between measurements and (estimated) state variables, and $\hat{e}$ is the mismatch accumulated in the process. The reason behind this convoluted approach is that the control center cannot access directly e, but $\hat{e}$ is computable. Thus, assuming to know the distribution of e and being able to deduce from that the distribution of $\hat{e}$, it is possible to define a random variable J that "summarizes" the information embedded in $\hat{e}$, i.e., a statistic that captures gross error. Then, to perform error detection it is possible to define a *test* to see if J is statistically relevant [8]. More formally, knowing the distribution of J and given a significance level α it is possible to compute τ, the α-quantile of J, and then design the single-tailed hypothesis test with H_0: $J \leq \tau$ and H_a: $J > \tau$. This is BSE. The test provides error detection, and related information provide also means to identify and remove bad measurements.

Summing up, the chain SE $\rightsquigarrow$ classic BSE $\rightsquigarrow$ bad measurements identification and clearance can tolerate at most $m - n$ *independent* faults (depending on system observability after measurements removal).

Notice that Σ can be considered a constant matrix over long periods of time, so τ can be computed once off-line and hard coded in the classic BSE. Figure 1 then shows only on-line computations, where the only variable seen from SE and BSE is z_a.

[2]Observability: if all state variables can be expressed as linear combinations of measurements, the state is *observable*; otherwise, it is not [7]. Observability can be determined investigating the null space of the rows of H.

[2]https://icseg.iti.illinois.edu/ieee-14-bus-system

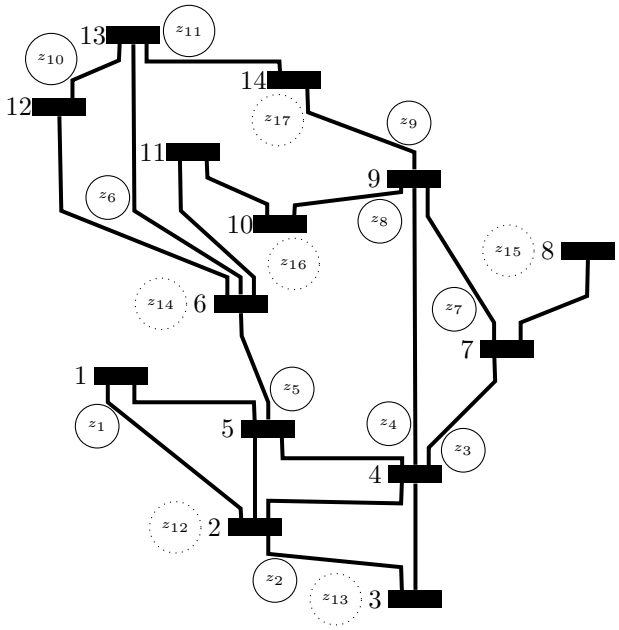


Fig. 2. Pictorial representation of the IEEE 14-bus[3] grid where "power injection" sensors are represented as dotted circles and "power flowing from" sensors are represented as full circles. In the example, $m = 17$ and $n = 13$. Generators and loads are not shown.

A. Classic (Bad) State Estimation for Linear Models

In the following, the *weighted least squares* approach, typically adopted in BSE related literature (e.g., in [1], [9]), is employed to obtain $\hat{x}$. Usually, it is assumed that the measurement error vector e is a zero-mean Gaussian vector with diagonal covariance matrix $\Sigma := \mathrm{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_m^2)$, $\sigma_i > 0$, call $W := \Sigma^{-1}$ the *weight matrix* and assume that h is linear[3], i.e., $h(x) = Hx$ with $H \in \mathbb{R}^{m \times n}$. It is easy to show that, calling $G := H^T W H$ the *gain matrix* and $E := G^{-1} H^T W$ the *estimator matrix*, the *estimate vector* $\hat{x} := Ez$ is the best[4] because it solves the weighted least

[3]If h is nonlinear, almost all the reasoning applies as well because of the Gauss-Newton algorithm (of course the involved random variable distributions are slightly different). This is well-known and studied, so it is not addressed here. For instance, see chapter 12.6 of [1], where in particular (12.30) defines $\hat{J}$ for nonlinear h, and [9]–[11] for more details.

[4]If h is linear then least squares estimator and maximum likelihood estimator coincide, so the "best" is not ambiguous. If h is nonlinear then things are slightly more complicated, but equally well studied.

square optimization problem:

$$\arg\min J(x), \text{ where}$$
$$J(x) := \big(z^T - x^T H^T\big) W \big(z - Hx\big) = e^T W e \sim \chi^2_m,$$

and χ^2_m is the chi-squared distribution with m degrees of freedom [8].

Under the assumption of independent sensor faults, it is possible to design the BSE as a statistical test. Unfortunately it is not possible to access directly the value of J because it is not possible to read e. This would have been ideal because J is χ^2_m distributed and defining statistical tests for J is easy [8]. Nevertheless, it is easy to show that $\hat{e}$ is equal to $\big(I - HG^{-1}H^T W\big)e$ and then to prove [5] that $\hat{e}$ is a zero-mean Gaussian vector with covariance matrix

$$\hat{\Sigma} = \mathbb{E}[\hat{e} \cdot \hat{e}^T] = \Sigma - HG^{-1}H^T,$$

that is a $m \times m$ symmetric matrix, but unfortunately has rank $m - n$, and then is not invertible. If $\hat{\Sigma}$ was invertible then $\hat{e}^T \hat{\Sigma}^{-1} \hat{e}$ would have been χ^2_m distributed. Two options are commonly considered to define the gross error statistic:

O1: define the random variable

$$\begin{aligned}\hat{J} :=& \hat{e}^T W \hat{e} = z^T \hat{W} z = \\ =& e^T \hat{W} e \sim \text{ generalized } \chi^2_{m-n}\end{aligned} \tag{1}$$

where $\hat{W} := W - WHG^{-1}H^T W$ and observe that $\hat{J}$ can be evaluated directly from z, so it is possible to perform BSE in parallel to SE, as depicted in Figure 1. Nowadays it is relatively easy to compute α-quartiles of this distribution [6], and then define the corresponding hypothesis test, but this choice has been rarely adopted in practice;

O2: observe that $\hat{J}$ is *almost*[7] χ^2_{m-n} distributed, and then it is possible to design the hypothesis test based on χ^2 distributions, that are much easier to manipulate. This choice is the most commonly adopted in practice because accidental hardware faults (the ones traditionally addressed) tend to present huge deviation of $\hat{J}$.

Notice that, even though it is not possible to evaluate the mismatch between estimate $\hat{x}$ and real values x, its covariance is known: $\mathbb{E}[\big(x - \hat{x}\big)\big(x^T - \hat{x}^T\big)] = G^{-1}$. Notice also that it is possible to define the normalized residual $\hat{e}^n_i := \hat{e}_i / \sqrt{\Sigma_{ii}}$ and observe [10], [11] that $\hat{e}^n_i$ has zero mean, unit variance, and is a Gaussian variable, so $\hat{J}^n := \big(\hat{e}^n\big)^T \cdot \hat{e}^n$ is χ^2_{m-n}-distributed.

[5]To show that the mean of $\hat{e}$ is zero, it is sufficient to observe that multiplying a Gaussian vector for a matrix defines a Gaussian vector with the same mean, and conclude observing that $\mathbb{E}[e] = 0$ by definition of e. For the covariance, one can exploit the fact that G is symmetric and $I - HG^{-1}H^T W$ is idempotent. This formula has been reported in (45) of [10].

[6]Indeed, only recently the function `gx2_params_norm_quad` of the MATLAB Generalized chi-squared toolbox made relatively easy to design such a test.

[7]For instance, a Kolmogorov-Smirnov test considering $m-n = 3$ degrees of freedom with $\alpha = 0.05$ and 30 samples can state that J and $\hat{J}$ are statistically the same (p-vale 0.6749), and this justifies the choice of many books (e.g., [1]) to not distinguish among them, but with 80 samples they are statistically relevant (p-value 0.0022).

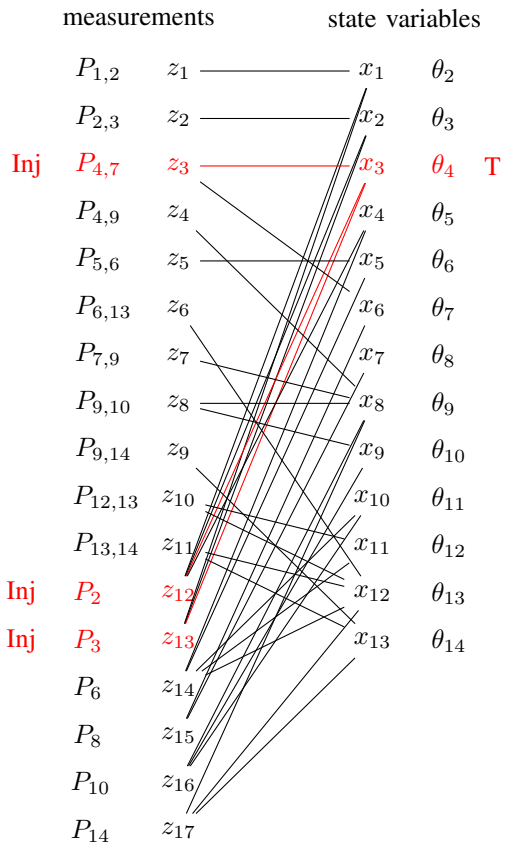


Fig. 3. Pictorial representation of the linear map h for the transmission grid of Figure 2: the nonzero entries of H define a bipartite (undirected) graph whose vertices are state variables and measurements. Power flowing from bus i to bus j is indicated $P_{i,j}$, power injected at bus i with P_i and voltage angle (with respect to the reference bus) at bus i with θ_i. As an example of FDI, the third state variable has been selected as the attack target (red T), so $c_i \neq 0$ iff $i = 3$, and correspondingly false data are injected (red Inj) into measurements linked to it, namely $a_i \neq 0$ iff $i \in \{3, 12, 13\}$.

Indeed, $\{\hat{e}^n_i\}_i$ can be exploited to identify those measurements that deviate too much (i.e., anomaly identification), but, being the focus of this paper on error/anomaly detection, no further detail is provided on this subject here.

B. False Data Injection

From the point of view of error handling in Dependability [2], there is (almost) no difference between accidental faults and intentional faults [8]: they are both faults. However, accidental faults are probabilistic whereas intentional are deterministic, and this difference in the nature of the fault has an impact on the efficacy of the techniques developed to cope with them. For the case of FDI attacks, unfortunately if the defender continues to apply the classic BSE, the consequences for the controlled infrastructure can be severe.

Consider the case [4] of an attacker that selects an *attack vector* a and adds it to (i.e., *injects it into*) the measure vector z, so the defender receives $z_a := z + a$ instead of z, as depicted

[8]Also called malicious faults, i.e., determined after an attack. The attack is moved from an insider or through an intrusion of an outsider.

in Figures 1 and 3. This means that the attacker intrudes the cyber part of sensors and alters the data that they transmit to the control center. The attacker's aim is to fool the defender in believing that the state is $x_a := x + c$ instead of x, where c is selected either at random (if the attacker's aim is just to force the defender at taking wrong decisions) or specifying certain values that the attacker knows will orient defender's decisions in a predetermined direction. Actually, the SE performed by the defender outputs $\hat{x}_a$ instead of $\hat{x}$, so the attacker wants $\hat{e}_a$ to be as close as possible to $\hat{e}$ for masking its actions. This is how FDI works.

In Section II it has been remarked that the classic BSE can tolerate independent (random) errors. The problem is that the classic BSE cannot tolerate *common-mode* errors, either determined by accidental faults (extremely rare) or by attacks (unfortunately plausible). In particular, if *no constraints on the attack are considered*, then the easiest choice for attackers is to set $a := Hc$ because $z_a - Hx_a = 0$, and then $\hat{J}_a$ and $\hat{J}$ are indistinguishable. More formally, if the attacker has:

A1: unlimited read access to the system information (i.e., h);
A2: write access to all the meters (i.e., al the entries of a can be nonzero);
A3: (if h is nonlinear) read access to all the meters and the ability to estimate the state (i.e., the attacker knows $\mathfrak{X}$ and $a(z, \hat{x})$ can be a function of measures and estimated states);

then FDI is easy. Of course such an attack is quite unrealistic, and there is a florid literature [12], [13] on FDI to address, from both attacker and defender points of view, the issue under several conditions, i.e., constraints on attacker knowledge or abilities. In particular, A3 is more involved [14] than the version reported here. The key point is that, to the best of the authors' knowledge, in all the reported attack/defense strategies the attackers have to know which information is exploited by the defender to perform BSE. Examples [4] of attacks (and corresponding constrains) are: address specific sensors (targeting specific state variables, with specific or random values), address a generic subset of sensors with a given cardinality (for specific or random values of given or generic state variables).

III. Random (Bad) State Estimator

One can design BSEs to perform all sort of checks on the measurements (e.g., seeing if the entries of z change magnitudes abruptly in few consecutive observations), to estimate states and also to integrate different sources of information. The statistical tests-based analysis of the classic BSE can then be part of a sophisticated component that performs error detection in a broad sense. The idea is: more information means more system awareness, and then better error detection. Apart from noticing that more information also means increased complexity, so designing SE and BSE can become challenging and costly, when intentional faults are addressed another important concept has to be considered: more information sources means larger *attack surface*, and then potentially less effective attack detection.

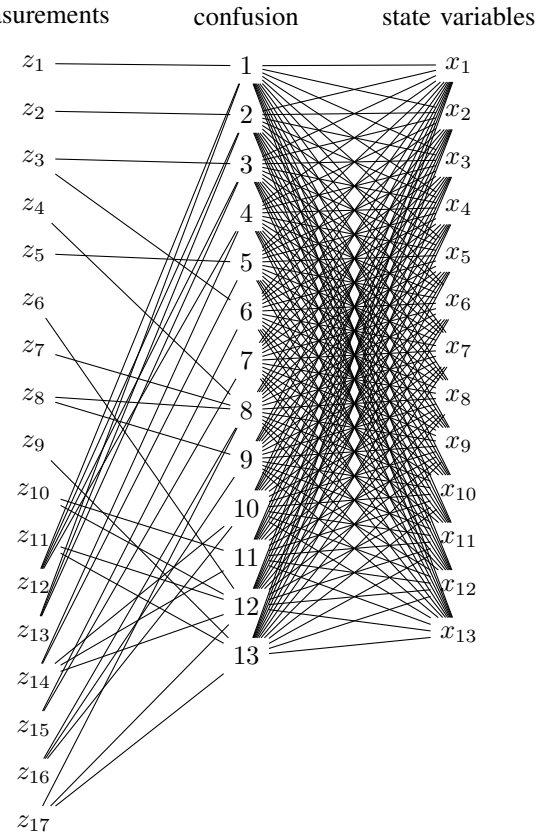


Fig. 4. Pictorial representation of confusion as a (random) additional layer between measurements and state variables. In particular, nonzero entries of M defines the confusion undirected graph.

For this reason, the design of R-BSE relies on the same information available to the classic BSE, just slightly increasing its complexity. The idea is to exploit the difference of information between defender and attackers: the defender knows $\mathfrak{X}$, while the attackers do not. In other words, assumption A3 of Section II-B is considered false, as usual in the context of linear BSE.

Moving from the observation that both attackers and classic BSE treat $H\hat{x}$ as a unique object with a physical meaning, namely $\hat{z}$, and this makes FDI possible, the idea at the heart of R-BSE is to decouple H and $\hat{x}$ introducing an additional layer of interactions between them, represented through the *confusion matrix* $M \in \mathbb{R}^{n \times n}$, as depicted in Figure 4. Clearly, the result has no more a physical meaning and it depends on the particular choice of M, that the defender can define at random with constraints detailed in Section III-A. This confounds the attacker, but makes more involved the definition of a gross error statistic that summarizes the mismatch resulted from going back and forth between measurements and state variables. Seen from a different perspective, deterministic (intentional) faults are now treated as random faults, where the source of randomness comes from the attacker's lack of knowledge of $\mathfrak{X}$ and M. Notice that R-BSE presents similarities with cryptographic schemes, but here no additional encoding or encryption of the

measurements z is involved with respect to the classic BSE.

More formally, define the *residue* vector as

$$r := z - HM\hat{x} = H(I - M)x + Be \neq \hat{z}, \tag{2}$$

where $B := I - HMG^{-1}H^TW$. The introduction of M makes r dependent on x, that is not accessible, so the main concern in designing R-BSE is how to make $H(I - M)x$ as close as possible to zero. In particular, if M is chosen so that $Mx = x$ then r is a zero-mean normal vector with covariance $B\Sigma B^T$, that is not invertible. Then

$$\tilde{J} := r^T r = e^T B^T Be \sim \text{ generalized } \chi^2_m \tag{3}$$

can be used as a statistic, and R-BSE designed as an hypothesis test. This approach is close to O1 in Section II-A. R-BSE is depicted in Figure 5. Notice that $\tilde{J}$ is different from the gross error $\hat{J}$, but still measures the deviation of the system from normal operation conditions.

Similarly to the link enforced by Equation (1) between the measurement error vector e and the classic BSE statistic $\hat{J}$, Equation (3) links e with the R-BSE statistic $\tilde{J}$. Thus, R-BSE is also able to address all the errors addressed by the classic BSE.

If the attacker knows M then choosing $a := HMc$ results in $r_a = r$, where $r_a := z_a - HM\hat{x}_a$, and then the FDI is as stealthy as in Section II-B. The point is that now M can be chosen by the defender at random, so the attacker cannot know M, despite the knowledge of all the system details, SE and BSE algorithms, unless the attacker can intrude the memory of the program running the algorithm depicted in Figure 5 or the seed exploited by the random number generator employed to update M (events that are usually excluded).

Notice that the standard FDI attack $a := Hc$ is no more stealthy because now $r_a = r + a - HMc \neq r$ and then $\tilde{J}_a \neq \tilde{J}$, so R-BSE can detect the error (when the classic BSE cannot), unless c is such that $Mc = c$ and then $a - HWc = 0$. The necessity of the attackers to know $\mathfrak{X}$ in order to define c is close to A3 of Section II-B, so R-BSE can be considered almost as difficult to fool as a nonlinear BSE.

Notice also that Equation (2) does not touch SE, so R-BSE can go in parallel with the classic BSE (that can remain in place), but it depends on $\hat{x}$, and then R-BSE needs to go in sequence with SE (see Appendix A for a proof of the impossibility to define a BSE that can simultaneously address FDI and operate independently from SE).

Is it magic? Actually, to work properly R-BSE requires assumptions on the state space, i.e., the defender has to know or approximate the proper subspace $\mathfrak{X}$ of $\mathbb{R}^n$ the state x belongs to between two system jumps, and this in some contexts can require too much information to be feasible, or even being impossible (when $\mathfrak{X} = \mathbb{R}^n$). Nonetheless, in those contexts where subsequent SEs produce slightly different values of x, e.g., electrical transmission grids, R-BSE is effective. Notice, though, that if the attacker knows $\mathfrak{X}$ (for instance by explicitly knowing a basis for it) then selecting c such that $Mc = c$ is easy. This eventuality is excluded because assumption A3 of Section II-B has been considered false throughout the paper.

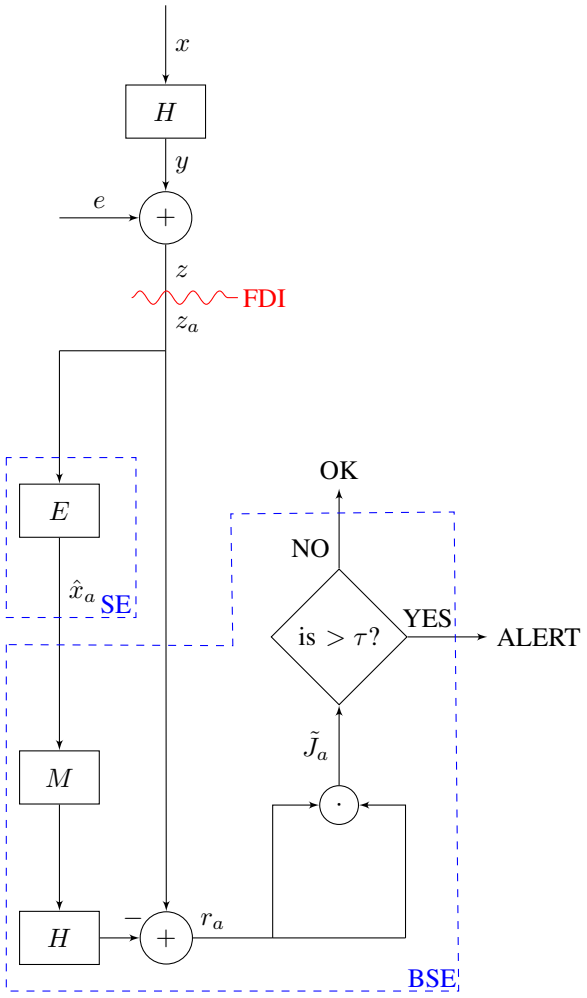


Fig. 5. Pictorial view of the on-line computations performed by R-BSE.

A. Definition of M, false positives and negatives

Assumed known an orthogonal basis $\{u_1, \ldots, u_k\}$ of $\mathfrak{X}$, there are several approaches to define M such that:

R1: M is orthogonal (i.e., it just "rotates" the space $\mathbb{R}^n$);
R2: M acts as the identity over $\mathfrak{X}$: $Mu = u, \forall u \in \mathfrak{X}$;
R3: M is "as random as possible" over $\mathfrak{X}^\perp$, the subspace of $\mathbb{R}^n$ orthogonal to $\mathfrak{X}$.

A simple way to accomplish the task is to complete the basis $\{u_1, \ldots, u_k\}$ by adding $n - k$ vectors $u_{k+1}, \ldots, u_n$ so that

$$M = V \begin{bmatrix} I_k & \\ & K \end{bmatrix} V^T, \qquad V = \begin{bmatrix} u_1 & \ldots & u_n \end{bmatrix}, \tag{4}$$

where $K \in \mathcal{O}(n - k)$, the Lie group of orthogonal matrices of size $n - k$.

To accomplish R3, K is prescribed to be drawn according to a probability measure μ over $\mathcal{O}(n-k)$, which is *uniform*: given any constant orthogonal matrix P, and Borelian measurable set $\mathcal{S}$, $\mu(P \cdot \mathcal{S}) = \mu(\mathcal{S})$ is required. Informally, "rotating" the group of orthogonal matrices by a constant action of the group should not modify the probability distribution. Note that this

choice in particular prescribes that reordering of the coordinates is irrelevant for the choice of K. Such measure is called left-invariant, and exists and is unique for compact Lie groups, for which it is called the Haar measure [15]. In practice, such a matrix can be easily sampled in MATLAB through a QR decomposition, by

```
>> [Q, R] = qr(randn(n-k));
>> K = Q / diag(diag(sign(R)));
```

The second step, normalizing the Q factor by the sign of the diagonal of R, forces the QR decomposition to have positive signs on R, which makes it unique and guarantees that the resulting K will be Haar distributed [15].

To build M in practice, a QR factorization of $[u_1, \dots, u_k]$ is computed, which gives a Q matrix with an orthogonal basis of $\mathfrak{X}$ in the first k columns and of $\mathfrak{X}^\perp$ in the remaining $n-k$, and the $Q = [U_1,\ U_2]$ is partitioned accordingly. Then, by defining

$$M := U_1 U_1^T + U_2 Z U_2^T, \qquad Z \sim \text{ Haar over } \mathcal{O}(n-k),$$

a matrix as in Equation (4) is obtained. It is worth noting that, if n is sufficiently large, then K can be sampled in an effective way by directly sampling its factorization in terms of Householder reflectors [16].

Notice that a large value of $k = \dim(\mathfrak{X})$ implies that the attacker has more chances to guess c such that $Mc = c$, i.e., higher probability of *false negative* (i.e., stealthy attack) in R-BSE. On the other hand, if $\mathfrak{X}$ changes (and consequently also M and τ change), but the defender's knowledge of such changes is not well aligned, then the probability of *false positive* increases because there are chances that $H(I-M)x$ in Equation (2) is not close to zero. Thus, the defender has to find trade-offs in designing the algorithm for updating its own knowledge about $\mathfrak{X}$, according to the criticality of the system. Studying such an algorithm is beyond the scope of this paper.

Nevertheless, given a generic vector v, it is possible to estimate how large is the norm of $H(I-M)v$, for a generic full column rank H. This provides, from one hand, an estimate of the probability of false negatives under the assumption that the defender's knowledge of $\mathfrak{X}$ is perfect, and, on the other hand, of false positives (to be added to α given by measurement errors) where there is no attack but the defender's knowledge of $\mathfrak{X}$ is not perfect. Note that if $v \in \mathfrak{X}$, then this norm is zero. Otherwise,

$$\|H(I-M)v\|_2 \geq \sigma_n(H)\|(I-M)v^\perp\|_2,$$

where $\sigma_n(H)$ is the n-th smallest singular value of H and $v = v^{\mathfrak{X}} + v^\perp$ with $v^{\mathfrak{X}} \in \mathfrak{X}$ and $v^\perp \in \mathfrak{X}^\perp$. Then

$$\|(I-M)v^\perp\|_2 = \|(I-Z)(U_2^T v^\perp)\|_2$$

To avoid false negatives, it is important to ensure that the event of having $(I-M)v^\perp$ small in norm is unlikely. Hence, a bound of the form

$$\mathbb{P}\{\|(I-M)v^\perp\|_2 \leq \alpha\} \leq F(\alpha)$$

is sought, with $F(\alpha) \to 0$ as $\alpha \to 0$. The norm of $(I-M)v^\perp$ is a random variable, and it is possible to assume that $w := U_2^T v^\perp$ is fixed; being Z invariant under multiplication by constant orthogonal matrices, it is not restrictive to consider an Householder reflector P such that $Pw = \|w\|_2 e_1$; hence,

$$\begin{aligned}\|(I-M)v^\perp\|_2 &= \|(I-Z)w\|_2 \sim \|(I-PZP)Pw\|_2 \\ &= \|(I-Z)e_1\|_2 \cdot \|w\|_2\end{aligned}$$

For $k = n-1$, Ze_1 is a discrete random variable with range $\{-1, 1\}$, each with probability 0.5. Therefore, this case will be discussed separately in Section IV. For $k < n-1$, the vector Ze_1 is a (normalized) Gaussian vector of length $n-k$ with components X_i with $i = 1, \dots, n-k$. Hence, ignoring its first component,

$$\|(I-Z)e_1\|_2^2 \geq \left(\sum_{i=1}^{n-k} X_i^2\right)^{-1} \sum_{i=2}^{n-k} X_i^2 = \frac{G}{X_1^2 + G}$$

where $G = \sum_{i=2}^{n-k} X_i^2 \sim \chi^2_{n-k-1}$, with G and X_1 independent. Note that, if $G = y$, it holds

$$\frac{y}{X_1^2 + y} \leq \alpha \iff |X_1| \geq y\left(\frac{1}{\alpha} - 1\right).$$

Since the PDF $g(y)$ of G is known, it follows that

$$\mathbb{P}\left\{\frac{G}{X_1^2+G} \leq \alpha\right\} = \int_0^\infty \mathbb{P}\left\{X_1^2 \geq y\left(\frac{1}{\alpha} - 1\right)\right\} g(y)dy.$$

Using the standard tail bound for normally distributed random variables

$$\mathbb{P}\{|X| \geq t\} \leq \frac{\sqrt{2}}{t\sqrt{\pi}} e^{-\frac{t^2}{2}},$$

and the PDF $g(y) = \frac{y^{\frac{n-k-3}{2}} e^{-\frac{y}{2}}}{\Gamma(\frac{n-k-1}{2}) 2^{\frac{n-k-1}{2}}}$, the integral for $k \leq n-3$ can be bounded by

$$\begin{aligned}\mathbb{P}\left\{\frac{G}{X_1^2+G} \leq \alpha\right\} &\leq \sqrt{2}\int_0^\infty \frac{e^{-\frac{y}{2}(\frac{1}{\alpha}-1)}}{\sqrt{\pi y(\frac{1}{\alpha}-1)}} g(y)dy \\ &= \frac{1}{\Gamma(\frac{n-k-1}{2})\sqrt{2^{n-k-1}\pi(\frac{1}{\alpha}-1)}} \int_0^\infty e^{-\frac{y}{2\alpha}} y^{\frac{n-k-4}{2}} dy \\ &= \frac{\Gamma(\frac{n-k-2}{2})}{\Gamma(\frac{n-k-1}{2})} \frac{1}{\sqrt{2\pi(\frac{1}{\alpha}-1)}} \alpha^{\frac{n-k-2}{2}}.\end{aligned}$$

Combining all these results, yields the sought bound

$$\mathbb{P}\{\|(I-M)v^\perp\|_2 \leq \alpha\|v^\perp\|_2\} \leq \frac{\alpha^{n-k-1}}{\sqrt{2}(1-\alpha)}, \quad \alpha < 1, \quad (5)$$

where we took advantage of $\Gamma(\frac{n-1}{2}) \leq \sqrt{\pi} \cdot \Gamma(\frac{n}{2})$ and of the property that $\Gamma(\frac{n}{2})$ is an increasing function for n integer and $n \geq 2$.

To be more concrete, given k and selecting $v := u_j$ with $k+1 \leq j \leq n$, the defender's knowledge imperfection can be modeled working with the state vector $x_0 + \epsilon \cdot v$, with $x_0 \in \mathfrak{X}$, for increasing values of $\epsilon \geq 0$.

B. Computational complexity

R-BSE comprises two groups of computations: update of M, as in Equation (4), and τ exploiting quantiles of the generalized χ^2_m distribution; and the evaluation of $\tilde{J}$ as in Equation (3). The former group is performed off-line and is required only when $\mathfrak{X}$ changes significantly, as a consequence of the system jumping from an equilibrium point to another, or periodically in order to increase attacker's confusion. This requires $\mathcal{O}((n-k)^3)$ floating point operations. The latter is performed on-line, every time new measurements are processed, and requires $\mathcal{O}(m)$ floating point multiplications. For large n, it may be worth considering the approach proposed in [16], which allows to build the matrix M in factored form in $\mathcal{O}((n-k)^2)$ flops, and to perform a matrix-vector product with M at the same quadratic complexity.

For comparison, the classic BSE needs no off-line update and the on-line evaluation of $\hat{J}$, performed in parallel to SE as depicted in Figure 1, requires $\mathcal{O}(m^2)$ operations. If the classic BSE is implemented in series to SE working with $\hat{e}$ and W, that is diagonal, it requires $\mathcal{O}(m)$ operations. The feasibility of R-BSE in the electrical transmission grid context has been analyzed in Section IV.

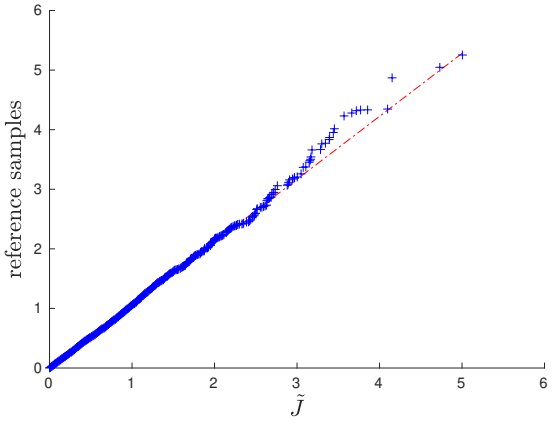
IV. Numerical evaluations

Numerical evaluations involved the IEEE 14-bus (depicted in Figure 2, where the attacked measurements are highlighted in red in Figure 3) and 300-bus[9] transmission grids. The computer where the experiments were performed has a 11^{th} Gen Intel(R) Core(TM) i7-1165G7 CPU, 4 unit clocked at 2.80GHz, 8 treads, 40Gb of DDR4 RAM clocked at 3200MHz. The operating system is Pop!_OS 20.04 LTS. The implementation of R-BSE has been written in Julia 1.5.3 and MATLAB, numerical experiments are performed exploiting JuliaGrid[10] and Generalized chi-square toolbox[11].

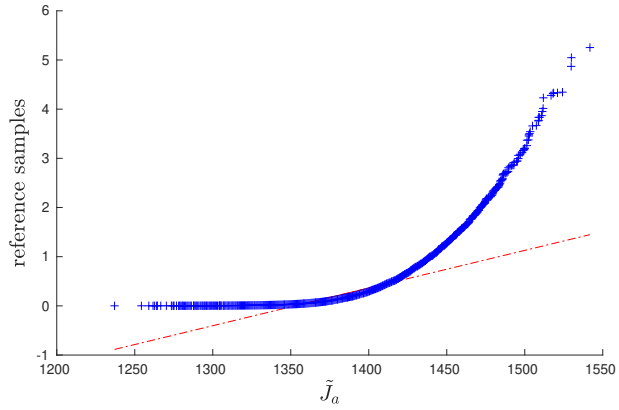
First of all, the time complexity reported in Section III-B has been checked and, for the 300-bus grid, the evaluation of $\tilde{J}$ with a dense M took an average of 4.85 milliseconds, and 0.24 milliseconds for $\hat{J}$. The off-line evaluation of τ took on average 11.01 seconds, with the minimum of 2 seconds.

Then, the fact that $\tilde{J}$ is (slightly less disperse than) generalized χ^2_m distributed, whereas $\tilde{J}_a$ is not, has been verified numerically for $k = 2$, as can be appreciated inspecting Figure 6. The quantile-quantile plots compare the empirical Cumulative Distribution Function (CDF) of $\tilde{J}$ and $\tilde{J}_a$ with the empirical CDF of samples drawn from the distribution determined in Equation (3). More formally, in Figure 6a the couples (q-quantile of $\tilde{J}$, q-quartile of generalized χ^2_m) are plotted for various values of q. Similarly, in Figure 6b the couples (q-quantile of $\tilde{J}_a$, q-quartile of generalized χ^2_m) are plotted. The closer these points are to the bisectrix, the better. Clearly, Figure 6a shows an almost perfect accordance, whereas Figure 6b shows a difference of about three orders of magnitude (representative of the case where attacks occur).

[9]https://icseg.iti.illinois.edu/ieee-300-bus-system
[10]https://github.com/mcosovic/JuliaGrid.jl
[11]https://github.com/abhranildas/gx2



(a) no attack



(b) under attack

Fig. 6. Quantile-quantile plot of $\tilde{J}$ for $k = 2$ in presence of no attack (a) and $\tilde{J}_a$ under attack (b) with respect to samples drawn from the χ^2_m distribution (10^3 samples) that are taken as a reference. As an example, the IEEE 14-bus grid depicted in Figure 2 and the attack illustrated in Figure 3 with $c_3 = 1$ have been considered.

Equation (5) has been verified for several values of k, an instance is shown in Figure 7, where the upper bound of the false negatives' CDF is checked. The proportion of false positives and negatives are shown in Figure 8. For this analysis, several values of k have been tested, for each of them several grid parameters assignments have been considered at increasing of ϵ. This way, in presence of no attack but assuming an increasingly imperfect knowledge of $\mathfrak{X}$, $||H(I-M)x||$ increases until $\tilde{J}$ overcomes the chosen threshold τ, thus producing a false positive. Similarly, in presence of attack with a defender that has perfect knowledge of $\mathfrak{X}$, if $\tilde{J}$ is not above τ then this is counted as a false negative.

As expected, for $\epsilon \leq 10^{-5}$ the proportion of false positives equals α, and increases at increasing of k. No false negative has been observed for $k < n-2$, and a few are observed for $k = n-2$, confirming that R-BSE detects a FDI of the kind described in Section II-B.

For $k = n-1$, the expression $(I-M)v$, with $v \in \mathfrak{X}^{\perp}$, can assume only two values: zero and $2v$, with probability 0.5.

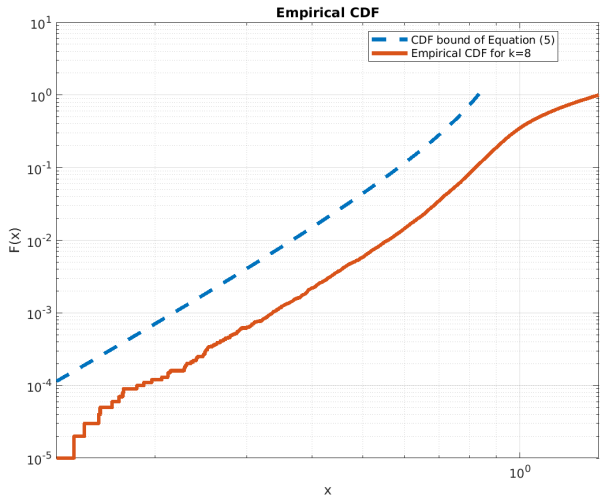


Fig. 7. Numerical check of Equation (5) for $\alpha = 0.05$ and $k = 8$ with 10^6 samples of M. Both x- and y- axes are in log scale.

Thus, according to the chosen value of τ, in presence of no measurement error the probability of false negatives is either zero or 0.5.

V. Related Work

The literature on BSE accounting for FDI has been mainly developed in the context of electrical grids, which is therefore taken as the reference for discussing related work. Key ideas on SE and BSE have been first reported in [10], [17], and the subject soon reached maturity (nowadays these concepts are common knowledge in electrical engineering [1], [11]). Relevant developments towards the kind of analysis presented in this paper are those in [18] on multiple errors, in [19] on common-mode errors, and a recent approach on hypothesis tests addressing attacks in [20].

Focusing on FDI, the seminal paper [4] systematized the subject and started a florid literature (an overview is provided in [12], [13]). The problem has been analyzed from different perspectives but, to the best of the authors' knowledge, no approach is directly comparable with R-BSE presented in this paper.

Several variations to the attack model presented in [4] have been addressed. Most of the studies assume that the attacker does not have access to complete information on the system structure and related parameters (such as [21]), or rely on only a subset of grid data, but collected over a period of time (such as [22]). Also the attack model in this paper assumes partial knowledge of the attacker, but explores a different direction than in previous studies, by assuming full knowledge of the grid structure and measurements, but partial knowledge of the state space that can be generated and needed to understand how the electrical grid moves from one equilibrium state to another one.

A different research direction concentrated on how new sensors (e.g., PMUs [23], [24] in addition to RTUs in the electrical context) can enhance not only SE but also BSE in addressing FDI.

Instead, R-BSE is sensor-agnostic, as soon as the physical model remains linear, because no special role is assigned to specific sensors. Moreover, we observe that, unavoidably, new equipment sooner or later becomes the target of attacks, and then attack surface and BSE design complexity increase.

Recently, in the literature the focus gradually shifted from linear to nonlinear model h [14], nowadays more commonly employed in several contexts. An example referred to FDI attack can be found in [25]. Differently from how R-BSE behaves, in these studies the assumption A3 discussed in Section II-B is usually enforced, and then detectors often rely on different mechanisms with respect to the classic BSE (for an example, see [26]).

Finally, R-BSE is tailored on dynamical systems close to equilibrium and stationary, so measurements useful for SE are taken with a frequency chosen to ignore transients. The very same data are exploited also by the classic BSE, and then by R-BSE. Other studies on FDI proposed anomaly detection based on data from different sources, e.g., transient data (collected periodically or after a stimulus [27]), or measurements on ICT components [28].

From the above reviewed literature, it is evident that the proposed R-BSE solution is not directly comparable with other approaches because of heterogeneity of assumptions or variety of information the specific techniques are based on. Therefore, in this paper we focused on assessing the enhancements of the new proposal with respect to the classic BSE, which is the reference baseline. A comparison study with other solutions requires careful planning to account for their inherent diversity, and it is therefore postponed to future work.

VI. Conclusions and Future Work

This paper presented R-BSE, an evolution of the classic and widely adopted BSE towards the management of FDI. Its strength consists in introducing a minimal number of changes to the classic solution, to essentially gain benefits in terms of costs and more straightforward applicability in contexts where the classic BSE is currently adopted. R-BSE leverages random elements in the analysis of measurements data to break the deterministic pattern of BSE, easily exploitable by an attacker to compromise carefully selected measurements in an undetectable manner. Application of the new techniques to representative IEEE transmission grid use cases show the enhancements of R-BSE with respect to the classic BSE.

Future work is foreseen in several directions. Among extensions of technical aspects to improve the efficacy of the proposed method, there is how to design an algorithm that approximates the vector space $\mathfrak{X}$. In addition, investigating more realistic formulations of $\mathfrak{X}$ than a vector space would bring benefits in lowering the chances an attacker may have to successfully accomplish the FDI attack.

Another interesting and timely research question to tackle in future work is how to adapt R-BSE to address nonlinear models h. It is expected that such adaptation process will follow a

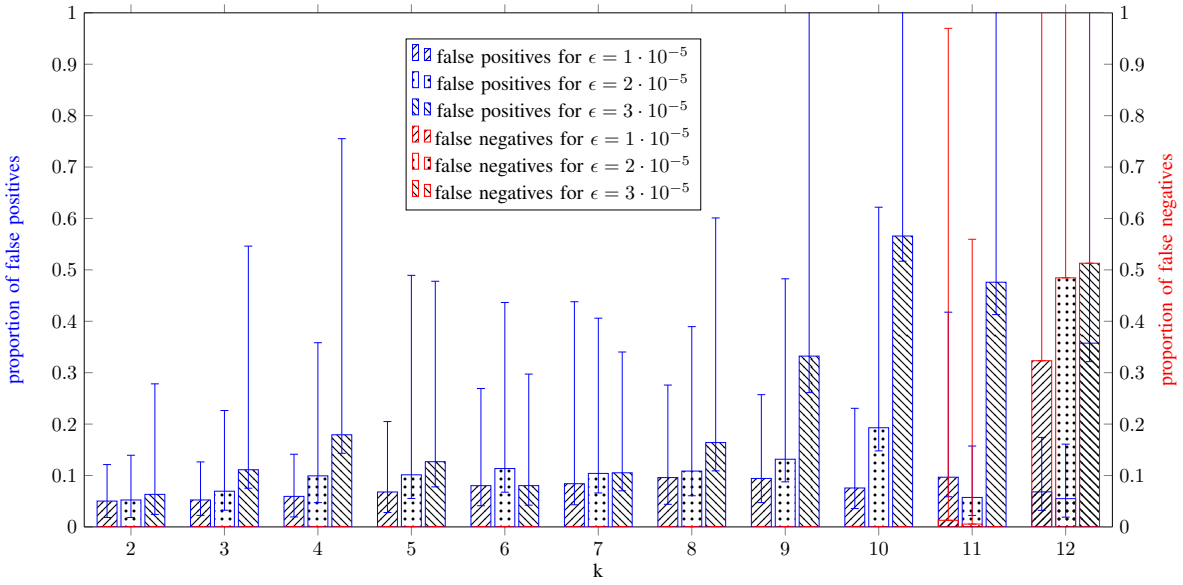


Fig. 8. Proportion of false positives (blue) and false negatives (red) over 100 samples of M, and for each of them 10^4 samples of e. The histogram reports the means, and for completeness also the min and max values are shown through segments on top of the histogram. The IEEE 14-bus grid depicted in Figure 2 and the attack described in Figure 3, where $a = Hc$ with $c_3 = 1$ and $c_i = 0$ for $i \neq 3$, have been taken as an example. The significance level in R-BSE is $\alpha = 0.05$.

similar path as taken by the classic BSE to nonlinear models, but the attack model may reveal peculiar differences.

Of course, extending the assessment campaign by adopting a variety of use cases and attack scenarios is another direction worth to elaborate more in depth.

Finally, developing a benchmark for conducting comparison with alternative solutions, accounting for the heterogeneity characterizing the different families, is a challenging but certainly useful development to pursue. Further, this kind of comparison would also pave the way to studies on identifying useful synergies among subset of anomaly detection techniques, so to attempt system designs that fruitfully compose them to improve attack detection, in accordance with security requirements of the application at hand.

APPENDIX

A. Issues with BSE in parallel with SE

Trying to design a statistic for BSE that is computable in parallel with the SE, and then relies only on z, produces a contradiction. Indeed, consider $\tilde{J} := z^T M z$, with M a generic matrix, so

$$\tilde{J} = x^T H^T M H x + x^T H^T \big(M + M^T\big) e + e^T M e.$$

At the same time, consider the standard FDI attack $a := Hc$, that produces

$$\tilde{J}_a = \tilde{J} + c^T H^T \big(M + M^T\big) z + c^T H^T M H e.$$

The issue is that, in order to make mean and covariance of $\tilde{J}$ computable, M has to be chosen so that $H^T M H = 0$ and $H^T \big(M + M^T\big) = 0$, but this implies $\tilde{J}_a = \tilde{J}$, so the attack is stealthy. Notice that the classic BSE is a special case of this, namely when $M := \hat{W}$.

ACKNOWLEDGMENT

We wish to thank Valerio Formicola from Siemens, USA, and Salvatore D'Antonio from the University Parthenope of Naples, Italy, for their fruitful discussion and insights on preliminary ideas at the basis of this work.

REFERENCES

[1] A. J. Wood, B. F. Wollenberg, and G. B. Sheblé, *Power Generation, Operation, and Control, 3rd Edition*, 2013.

[2] A. Avizienis, J.-C. Laprie, B. Randell, and C. Landwehr, "Basic concepts and taxonomy of dependable and secure computing," *IEEE Transactions on Dependable and Secure Computing*, vol. 1, no. 1, pp. 11–33, 2004.

[3] D.-T. Peng, J. Dong, and Q. Peng, "Overloaded branch chains induced by false data injection attack in smart grid," *IEEE Signal Processing Letters*, vol. 27, pp. 426–430, 2020.

[4] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Transactions Information and System Security*, vol. 14, no. 1, 2011.

[5] V. J. Hodge and J. Austin, "A survey of outlier detection methodologies," *Artificial Intelligence Review*, vol. 22, pp. 85–126, 2004.

[6] V. Barnett and T. Lewis, *Outliers in statistical data*, ser. Wiley Series in Probability and Mathematical Statistics. Applied Probability and Statistics, 1984.

[7] E. Castillo, A. J. Conejo, R. E. Pruneda, and C. Solares, "Observability analysis in state estimation: a unified numerical approach," *IEEE Transactions on Power Systems*, vol. 21, no. 2, pp. 877–886, 2006.

[8] A. M. Mood, "Introduction to the theory of statistics," 1950.

[9] F. C. Schweppe and J. Wildes, "Power system static-state estimation, part i: Exact model," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-89, no. 1, pp. 120–125, 1970.

[10] F. C. Schweppe and D. B. Rom, "Power system static-state estimation, part ii: Approximate model," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-89, no. 1, pp. 125–130, 1970.

[11] A. Monticelli, "Electric power system state estimation," *Proceedings of the IEEE*, vol. 88, no. 2, pp. 262–282, 2000.

[12] R. Deng, G. Xiao, R. Lu, H. Liang, and A. V. Vasilakos, "False data injection on state estimation in power systems—attacks, impacts, and defense: A survey," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 2, pp. 411–423, 2017.

[13] S. Aoufi, A. Derhab, and M. Guerroumi, "Survey of false data injection in smart power grid: Attacks, countermeasures and challenges," *Journal of Information Security and Applications*, vol. 54, p. 102518, 2020.
[14] G. Hug and J. A. Giampapa, "Vulnerability assessment of ac state estimation with respect to false data injection cyber-attacks," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1362–1370, 2012.
[15] A. Edelman and N. R. Rao, "Random matrix theory," *Acta numerica*, vol. 14, pp. 233–297, 2005.
[16] G. W. Stewart, "The efficient generation of random orthogonal matrices with an application to condition estimators," *SIAM Journal on Numerical Analysis*, vol. 17, no. 3, pp. 403–409, 1980.
[17] R. E. Larson, W. F. Tinney, and J. Peschon, "State estimation in power systems part i: Theory and feasibility," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-89, no. 3, pp. 345–352, 1970.
[18] K. A. Clements and P. W. Davis, "Multiple bad data detectability and identifiability: A geometric approach," *IEEE Transactions on Power Delivery*, vol. 1, no. 3, pp. 355–360, 1986.
[19] T. V. Cutsem, M. Ribbens-Pavella, and L. Mili, "Hypothesis testing identification: A new method for bad data analysis in power system state estimation," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-103, no. 11, pp. 3239–3252, 1984.
[20] A. S. Bretas, N. G. Bretas, B. Carvalho, E. Baeyens, and P. P. Khargonekar, "Smart grids cyber-physical security as a malicious data attack: An innovation approach," *Electric Power Systems Research*, vol. 149, pp. 210–219, 2017.
[21] M. A. Rahman and H. Mohsenian-Rad, "False data injection attacks with incomplete information against smart power grids," in *IEEE Global Communications Conference (GLOBECOM)*, 2012, pp. 3153–3158.
[22] S. Lakshminarayana, A. Kammoun, M. Debbah, and H. V. Poor, "Data-driven false data injection attacks against power grids: A random matrix approach," *IEEE Transactions on Smart Grid*, vol. 12, no. 1, pp. 635–646, 2021.
[23] S. Basumallik, S. Eftekharnejad, N. Davis, and B. K. Johnson, "Impact of false data injection attacks on pmu-based state estimation," in *North American Power Symposium (NAPS)*, 2017, pp. 1–6.
[24] S. Almasabi, T. Alsuwian, E. Javed, M. Irfan, M. Jalalah, B. Aljafari, and F. A. Harraz, "A novel technique to detect false data injection attacks on phasor measurement units," *Sensors*, vol. 21, no. 17, 2021.
[25] C. Konstantinou and M. Maniatakos, "A case study on implementing false data injection attacks against nonlinear state estimation," in *Proceedings of the second ACM workshop on cyber-physical systems security & privacy*, 2016.
[26] J. J. Q. Yu, Y. Hou, and V. O. K. Li, "Online false data injection attack detection with wavelet transform and deep neural networks," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 7, pp. 3271–3280, 2018.
[27] K. R. James Ranjith, D. Kundur, and B. Sikdar, "Transient model-based detection scheme for false data injection attacks in microgrids," in *IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, 2019, pp. 1–6.
[28] W. Yu, D. Griffith, L. Ge, S. Bhattarai, and N. Golmie, "An integrated detection system against false data injection attacks in the smart grid," *Security and Communication Networks*, vol. 8, no. 2, pp. 91–109, 2015.